



UNIVERSITÀ DEGLI STUDI DI MACERATA

**CORSO DI DOTTORATO DI RICERCA IN
UMANESIMO E TECNOLOGIE**

CICLO XXXVI

TITOLO DELLA TESI

Estrazione della conoscenza da immagini nel dominio del patrimonio culturale

SUPERVISORI

Chiar.ma Prof. Anna Rovella

Chiar.mo Prof. Stefano Pigliapoco

DOTTORANDO

Dott. Martin Critelli

COORDINATORE

Chiar.mo Prof. Roberto Lambertini

ANNO 2024



Indice

Lista degli Acronimi	4
1. Introduzione	8
2. Il Patrimonio culturale	19
Patrimonio culturale, identità dei popoli e attrattività turistica dei territori.....	28
Il processo di digitalizzazione del patrimonio culturale in Italia.	29
3. Le metodologie di descrizione standardizzata di immagini come strumento di diffusione e valorizzazione del patrimonio culturale.....	37
Il modello International Image Interoperability Framework	37
Funzionamento del IIIF.....	38
La descrizione delle risorse secondo lo standard IIIF.....	46
Casi applicativi.....	51
4. Analisi delle immagini e riconoscimento automatico di immagini statiche dinamiche e audiovideo: metodologie e strumenti.....	61
Immagini statiche e video	63
Reti Neurali e riconoscimento di testi manoscritti digitalizzati.....	74
Audio e trattamento di audiovideo.....	76
5. Lo standard IIIF per la valorizzazione della conoscenza estratta automaticamente tramite Reti Neurali Convolutionali.....	94
6. Prospettive future e conclusioni.....	101
Bibliografia	122
APPENDICE 1 – Scheda descrittiva in VRA Core de La Scuola di Atene.....	142
APPENDICE 2 – Esempio del file JSON-LD per il modello di <i>Collection</i> IIIF.....	144
APPENDICE 3 – Risultati video activity recognition.....	147
APPENDICE 4 – Estrazione automatica delle informazioni dai testi	151

Lista degli Acronimi

ACS	Archivio Centrale dello Stato
AGID	Agenzia per l'Italia Digitale
AI	Artificial Intelligence
AIFF	Audio Interchange File Format
AIFF-C	Audio Interchange File Format Compressed
API	Application Profile Interface
AR	Augmentaty Reality
ASR	Automatic Speech Recognition
AURORA	Artwork Unique Recognition and Tracking
AWI TS	Approved Work Item Technical Subcommittee
BDD100K	Berkeley DeepDrive
BERT	Bidirectional Encoder Representations from Transformers
CAD	Codice dell'Amministrazione Digitale
CDWA	Categories for the Description of Works of Art
CEN	Comitato europeo di normalizzazione
CENELEC	Comitato europeo di normalizzazione elettronica
CH	Cultural Heritage
CIDOC	International Committee for Documentation
CIDOC CRM	International Committee for Documentation Conceptual Reference Model
CNNs	Convolutional Neural Networks
CNR-ILC	Consiglio Nazionale delle Ricerche - Istituto di Linguistica Computazionale
COMEST	UNESCO World Commission on Ethics of Scientific Knowledge and Technology
COVID-19	Coronavirus Disease
CTC	Connection Temporal Classification
CULT	Committee on Culture and Education
D.Lgs	Decreto Legislativo
DC	Dublin Core
DCMI	Dublin Core Metadata Initiative
DH	Digital Humanities

DL	Deep Learning
DM	Decreto Ministeriale
DMP	Data Management Plan
DVL	Digital Vatican Library
EAD	Encoded Archival Description
EAP	Engagered Archives Programme
EDL	Estense Digital Library
FAIR	Findable Accessible Interoperable Resuable
FAM	Europeana Aggregators' Forum
FCL	Fully Connected Layer
FCS	UNESCO Framework for Cultural Statistics
FITS	File Information Tool Set
FLAC	Free Lossless Audio Codec
FLOs	Floating-point operations
GDPR	General Data Protection Regulation
GLAM	Galleries, Libraries, Archives and Museums
HMM	Hidden Markov Model
HTTP	Hypertext Transfer Protocol
ICCD	Istituto Centrale per il Catalogo e la Documentazione
ICCU	Istituto Centrale per il Catalogo Unico
ICH	Intangible Cultural Heritage
ICT	Information and Communication Technologies
IEC	International electrotechnical commission
IFLA	International Federation of Library Associations and Institutions
IIIF	International Image Interoperability Framework
ILSVRC	Large Scale Visual Recognition Challenge
ISAD (G)	General International Standard Archival Description
ISBD	International Standard Bibliographic Description
ISNART	Istituto Nazionale Ricerche Turistiche
ISO	International Organization for Standardization
JSON-LD	JavaScript Object Notation for Linked Data

JTC	Joint Technical Committee
KITTI	Karlsruhe Institute of Technology and Toyota Technological Institute
LoC	Library of Congress
LOD	Linked Open Data
LPC	Linear Predictive Coding
LSTM	Long-Short Term Memory
MAB	Musei Archivi e Biblioteche
MARC	MAchine Readable Cataloguing
METS	Metadata Encoding and Trasmission Standard
MFCC	Mel-frequency cepstrum
MiC	Ministero della Cultura
ML	Machine Learning
MLOL	Media Library OnLine
MLOL-DH	Media Library OnLine - Digital Humanities
MODS	Metadata Object Description Schema
MPEG3	Motion Pictures Coding Experts Group
MS COCO	Microsoft Common Objects in Context
NER	Named Entity Recognition
NISO MIX	National Information Standards Organization Metadata for Images In XML Standard
NLTK	Natural Language ToolKit
OAI-PMH	Open Archives Initiative Protocol for Metadata Harvesting
OCR	Optical Character Recognition
ONU	Organizzazione delle Nazioni Unite
PASCAL VOC	Pattern Analysis, Statical Modelling and Computational Learning Visual Object Classes
PND	Piano nazionale di digitalizzazione del patrimonio culturale
PNRR	Piano nazionale di ripresa e resilienza
POS	Part of Speech Tagging
PREMIS	Preservation Metadata: Implementation Strategies
RASTA	RelAtive SpecTraiAL
R-CNN	Region-Based CNN
RGB	Red Green Blue

RNNs	Recurrent Neural Networks
RoI	Region of Interest
SVM	Support Vector Machine
TINT	The Italian NLP Tool
UCF 101	University of Central Florida Action Recognition Data Set
UE	Unione Europea
UNESCO	United Nations for Education, Science and Culture
URI	Uniform Resource Identifier
VCR	Visual Commonsense Reasoning
VeDPH	Venice Centre for Digital and Public Humanities
VIAF	Virtual International Authority File
VR	Virtual Reality
VRA	Visual Resources Association
W3C	World Wide Web Consortium
WAV	Waveform Audio File Format
XML	eXtensible Markup Language
XSD	XML Schema Definitions
YOLO	You Only Look Once

1. Introduzione

L'aspetto etico nei sistemi di intelligenza artificiale è un elemento che deve interessare non soltanto i processi di implementazione ma anche le modalità e gli obiettivi attraverso i quali essi vengono applicati¹. Occorre, dunque, comprendere in che modo le tecnologie influenzano i rapporti tra società, economia e cultura² e impattano sui processi che consentono la tutela e la valorizzazione del patrimonio culturale. Le sfide poste in essere sono, per loro natura, contrassegnate di un fondamento etico imprescindibile³. È questo un tema su cui si discute da decenni ma che solo negli ultimi anni è entrato nel merito dei rischi e delle opportunità connesse all'uso dell'AI per le società⁴ che ha avuto come effetto la definizione di principi etici per un uso corretto dell'intelligenza artificiale. Il rischio, come sottolineato da Luciano Floridi e Josh Cows⁵, è quello di una ridondanza e sovrapposizione di principi, talvolta simili, che finiscono col creare ambiguità in un contesto complesso di per sé. In risposta al proliferare di principi, si inserisce il framework, proposto dai due autori qui citati, per definire un quadro comune di valutazione dell'AI individuando i seguenti principi:

- **Beneficence**, comprende le attività volte a promuovere il “*well-being*” e di cui l'umanità possa godere di un beneficio.
- **Non-maleficence**, con l'espressione “non fare del male” sono intesi i principi volti a tutelare la sfera privata dell'individuo e a promuovere un uso quanto più sicuro possibile dell'AI.
- **Autonomy**, si riferisce al potere decisionale che gli individui trasferiscono verso i sistemi di intelligenza artificiale. Il discorso entra, dunque, nel merito del grado di accettabilità del rischio connesso con l'autonomia decisionale dei sistemi a discapito di quella umana.
- **Justice**, si tratta di un principio fortemente dipendente dal contesto e, come tale, difficilmente valutabile a priori.
- **Explicability**, attiene alla comprensione e alla trasparenza dei sistemi di intelligenza artificiale inserendosi come elemento complementare per la valutazione degli altri principi.

¹ Irina Aseeva, and Vladimir Budanov, “Digitalization: potential risks for civil society,” *Economic Annals – XXI* 186, n. 1 (2020). DOI: 10.21003/ea.V186-05.

² Luciana Lazzaretti, “What is the role of culture facing the digital revolution challenge? Some reflections for a research agenda,” *European Planning Studies* 30, n. 9 (2022). DOI: 10.1080/09654313.2020.1836133

³ Luciano Floridi, “Translating Principles into Practices of Digital Ethics: Five Risks of Being Unethical,” *Philosophy & Technology* 32, (2019). DOI: 10.1007/s13347-019-00354-x.

⁴ Luciano Floridi, and Josh Cows, “A Unified Framework of Five Principles for Ethical AI,” *Harvard Data Science Science Review* 1, n. 1 (2019). DOI: 10.1162/99608f92.8cd550d1.

⁵ Luciano Floridi, and Josh Cows, “A Unified Framework of Five Principles for Ethical AI,” *cit.*, p. 4.

Insieme alle conseguenze di carattere etico poste in gioco sul piano sociale⁶, l'applicazione delle tecnologie di intelligenza artificiale impatta anche sull'aspetto pragmatico e le modalità attraverso cui ciò avviene. Se da un lato, le tecnologie costituiscono un prezioso medium per apportare benefici, occorre sottolineare anche gli aspetti negativi legati ad un loro uso improprio, e, talvolta, illegale, che può avere gravi ripercussioni sugli individui e i territori. In tale scenario le tecnologie possono rappresentare una minaccia per il patrimonio culturale che, calato in nuovo contesto, si trova a dover fronteggiare pericoli occorsi, storicamente, in circostanze e con modalità tangibili. È quantomai importante prendere piena consapevolezza delle responsabilità che lo sviluppo e l'uso di tali tecnologie comportano. Nel contesto preso in esame nel presente lavoro di dottorato, l'applicazione dell'intelligenza artificiale permette certamente di delineare un quadro ampio aprendo a nuove possibilità e soluzioni. Come indicato anche nel Report “*AI in Education, Culture and the Audiovisual sector*”⁷ approvato dal Parlamento Europeo nell'aprile 2021, le potenzialità applicative in un dominio così ampio e complesso come quello del patrimonio culturale, possono essere molteplici. Il Report, stilato a seguito dell'udienza tenuta nel febbraio del 2020 dal Committee on Culture and Education (CULT Committee) sul tema “*The use of Artificial Intelligence in education, culture and the audiovisual sector*” ha l'obiettivo di analizzare l'impatto attuale e futuro dell'AI in settori della conoscenza e della comunicazione strategici quali la cultura, l'educazione e l'audiovideo. Il documento raccoglie le varie osservazioni emerse durante l'udienza classificandole tra generali e specifiche. Le prime focalizzano l'attenzione sull'importanza strategica assunta dall'intelligenza artificiale per l'Unione Europea nei diversi contesti disciplinari sottolineando la centralità dell'uomo nel loro sviluppo e impiego. L'accento è dunque posto sui diritti dell'uomo e su un uso etico delle tecnologie affinché esse siano «strumento al servizio delle persone, del bene comune e dell'interesse generale dei cittadini»⁸. Le osservazioni generali delineano un uso dell'AI nel rispetto della Carta dei Diritti Fondamentali dell'Unione Europea⁹ ribadendo la necessità che «gli algoritmi e l'IA dovrebbero essere “etici fin dalla loro progettazione”, senza pregiudizi incorporati»¹⁰. Il pericolo qui avvertito è legato ad un uso spregiudicato e discriminatorio delle tecnologie che comporterebbe, inevitabilmente, distorsioni

⁶ Anna V. Guryanova et al., “Socio-ethical Problems of the Digital Economy: Challenges and Risks,” *Digital Transformation of the Economy: Challenges, Trends and New Opportunities. Advances in Intelligent Systems and Computing* 908. DOI: 10.1007/978-3-030-11367-4_9.

⁷ Committee on Culture and Education, *AI in Education, Culture and the Audiovisual sector* (Parlamento Europeo 2021), https://www.europarl.europa.eu/doceo/document/TA-9-2021-0238_IT.pdf.

⁸ Committee on Culture and Education, *AI in Education, Culture and the Audiovisual sector, cit.*, punto 1, https://www.europarl.europa.eu/doceo/document/TA-9-2021-0238_IT.pdf.

⁹ Commissione Europea, 2016/C 202/389 *Charter of Fundamental Rights of the European Union* (2016), <https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:12016P/TXT>.

¹⁰ Committee on Culture and Education, *AI in Education, Culture and the Audiovisual sector, cit.*, punto 5, https://www.europarl.europa.eu/doceo/document/TA-9-2021-0238_IT.pdf.

cognitive e disuguaglianze di ogni genere. È in questo scenario che si manifesta l'importanza di produrre strumenti di AI affidabili, specie nel campo dell'istruzione, della cultura e dell'audiovisivo «per correggere stereotipi di genere e le distorsioni di genere, che possono avere un impatto negativo»¹¹. Anche la sicurezza dei dati personali è tra i temi interessati dalle osservazioni¹² con un'attenzione speciale verso la tutela dei dati dei minori. In linea generica, le osservazioni ribadiscono il rispetto del GDPR (*General Data Protection Regulation*) come vincolo necessario per la diffusione dei sistemi di intelligenza artificiale. Il rapporto tra dati personali e intelligenza artificiale è un tema estremamente delicato su cui anche il Garante della privacy è intervenuto con una serie di video¹³, volti a informare i cittadini sul tema, e raccogliendo la documentazione prodotta a riguardo dalle altre autorità quali, ad esempio, l'interim report prodotto dall'ONU "*Governing AI for Humanity*"¹⁴. Le osservazioni di carattere generale implicano anche la cybersicurezza e le potenzialità, offerte dall'AI, in termini di lotta alla criminalità informatica nel rispetto sempre dei diritti fondamentali. Quest'ultimo passaggio sottolinea la necessità di rendere le soluzioni basate su AI resilienti ad ogni genere di attacco informatico al fine di non compromettere la sicurezza e i dati ad esse affidate.

In seguito alle osservazioni di carattere generale, il documento riporta anche quelle specifiche relative ai settori presi ad esame: istruzione, cultura – e dei creatori di contenuti culturali – e audiovisivo.

Per ciò che concerne il settore dell'istruzione, il Report evidenzia la determinazione della Comunità europea a rafforzare le competenze digitali e garantire un maggiore livello di alfabetizzazione informatica (*digital literacy*) ai cittadini. Le tecnologie di intelligenza artificiale devono costituire, a tale scopo, uno strumento a disposizione delle scuole a garantire un'istruzione di qualità contribuendo a ridurre il *digital gap*¹⁵. Occorre prestare molta attenzione ad assicurare un accesso paritario alle tecnologie per evitare che queste diventino, contrariamente alle aspettative, motivo di una maggiore disparità sociale e territoriale. Parimenti, non è possibile considerare la sola disponibilità delle risorse sufficiente a colmare il divario nell'alfabetizzazione informatica se, in aggiunta, non si dispongono delle conoscenze necessarie a sfruttare adeguatamente le tecnologie. Per il ruolo che rivestono, è infatti indispensabile che gli insegnanti siano coinvolti dalle comunità di sviluppatori affinché «siano

¹¹ Committee on Culture and Education, *AI in Education, Culture and the Audiovisual sector*, cit., punto 11, https://www.europarl.europa.eu/doceo/document/TA-9-2021-0238_IT.pdf.

¹² Committee on Culture and Education, *AI in Education, Culture and the Audiovisual sector*, cit., punto 16, https://www.europarl.europa.eu/doceo/document/TA-9-2021-0238_IT.pdf.

¹³ Garante per la protezione dei dati personali, "AI – Intelligenza artificiale," <https://www.garanteprivacy.it/temi/intelligenza-artificiale>. Ultimo accesso 7 Gennaio 2024.

¹⁴ United Nations Organization (ONU), *Interim Report: Governing AI for Humanity* (2023), <https://www.un.org/ai-advisory-body>.

¹⁵ Angelo Alù, and Alessandro Longo, "Cos'è il digital divide, nuova discriminazione sociale (e culturale)," *Agenda Digitale*, 13 Marzo 2020, <https://www.agendadigitale.eu/cultura-digitale/il-digital-divide-culturale-e-una-nuova-discriminazione-sociale/>.

in grado di adattarsi a contesti d'istruzione basati sull'IA e di acquisire le conoscenze e le competenze necessarie per utilizzare le tecnologie di IA in modo pedagogico e costruttivo»¹⁶. Anche il ruolo assunto dei *player* tecnologici privati nel settore dell'istruzione è oggetto di indagine nel Report, per i quali si evidenzia, senza non poche preoccupazioni, il potere attrattivo su ricercatori e studiosi rispetto alle istituzioni pubbliche e alla ricerca accademica, insieme al crescente monopolio nella fornitura di servizi tecnologici per l'istruzione. Su quest'ultimo punto il Parlamento Europeo sostiene la necessità di investire adeguati finanziamenti a far sì che anche le piccole e medie imprese (PMI) possano avere un ruolo attivo nel mercato secondo le regole europee di concorrenza.

In una società abituata a esistere sempre di più in rete e nella quale i sistemi di comunicazione e interazione sono condizionati dai social network, gli audiovideo costituiscono un medium largamente adottato dalle istituzioni per la veicolazione di notizie, eventi e comunicazioni di vario genere. L'applicazione dell'intelligenza artificiale per lo sviluppo di algoritmi decisionali in grado di suggerire e mostrare mirati contenuti agli utenti costituisce un fattore ad alto rischio per la società. Come tale, si sottolinea nel Report la richiesta da parte del Parlamento Europeo che «gli algoritmi di raccomandazione e la pubblicità personalizzata siano spiegabili e trasparenti nella misura del possibile, così da consentire ai consumatori di comprendere in modo preciso e completo tali processi e contenuti»¹⁷. È essenziale che il consumatore sia consapevole di che ciò che gli viene mostrato, tutelando il diritto a non acconsentire ai servizi di personalizzazione o raccomandazione. Insieme ai servizi di promozione, anche la creazione di contenuti digitali è essere oggetto di trattazione nelle osservazioni in merito gli audiovideo, rafforzando, rispetto a quanto già previsto tra le considerazioni generali, il rispetto delle diversità culturali e di genere per evitare una qualsivoglia forma di discriminazione. Tra i maggiori pericoli connessi con un uso improprio delle tecnologie per la creazione di contenuti digitali rientra senza dubbio alcuno la disinformazione. Il *deepfake* è esempio di una prassi sempre più popolare di applicazione, a vario titolo, dell'intelligenza artificiale per la creazione di contenuti falsi dai quali diviene sempre più difficile discenderne la verità. Al giorno d'oggi, la creazione di tal genere di contenuti assume un ruolo considerevole per il settore creativo, basti pensare alle decine di app sviluppate in merito o alle funzionalità su alcuni social network che consentono a chiunque di “giocare” tra realtà e finzione. Se da una parte si rileva dunque un impatto positivo di tali attività sui settori culturali, e non solo, dall'altra emerge il pericolo che esse possano essere applicate con modalità illecite al fine di arrecare danni, di qualunque genere, a individui e

¹⁶ Committee on Culture and Education, *AI in Education, Culture and the Audiovisual sector, cit.*, punto 36, https://www.europarl.europa.eu/doceo/document/TA-9-2021-0238_IT.pdf.

¹⁷ Committee on Culture and Education, *AI in Education, Culture and the Audiovisual sector, cit.*, punto 77, https://www.europarl.europa.eu/doceo/document/TA-9-2021-0238_IT.pdf.

società. È quanto mai necessario regolamentare e stabilire un quadro giuridico entro il quale pratiche del genere debbano essere limitate¹⁸.

Relativamente al settore culturale, nelle osservazioni viene enfatizzato l'apporto positivo che le tecnologie di AI possono offrire e, in particolare, nel contrasto degli effetti negativi a seguito di fenomeni ostili quali: l'erosione e il deterioramento dei materiali, la distruzione per catastrofi naturali o per conflitti, l'individuazione di traffici illeciti di vario tipo. A tal proposito, nel Report sono anche identificate, con il contributo delle diverse Commissioni che compongono il Comitato (*Committee of Cultural and Education, Committee on Civil Liberties, Justice and Home Affairs, Committee on the International Market and Consumer Protection, Committee on Legal Affairs e, infine, Committee on Women's Rights and Gender Equality*), modalità e casi esemplificativi attraverso cui l'impiego dell'AI permette la salvaguardia e promozione del patrimonio culturale. Si distingue fra diverse sfere di interesse in cui l'uso dell'intelligenza artificiale sul patrimonio culturale può apportare dei benefici. In materia di accessibilità, l'uso del 3D e della realtà aumentata consente una fruizione sempre più estesa dei beni, soprattutto in termini turistici, con la possibilità di personalizzare i contenuti per persona, introdurre nuovi servizi interattivi e agevolare l'esperienza a persone portatori di disabilità. Anche in termini di ricchezza semantica le tecnologie possono avere un forte impatto offrendo nuovi strumenti e metodologie per la creazione di banche dati uniformi e sistemi di classificazione funzionali a migliorare la descrizione degli oggetti culturali ponendo le basi per la realizzazione di collegamenti tra gli stessi. Sul piano della sicurezza, si sottolinea l'importanza dell'AI per l'implementazione di strumenti destinati a combattere i traffici illeciti e la distruzione dei beni culturali¹⁹. Si tratta di fenomeni che raggiungono ogni anno ingenti numeri²⁰ contro i quali l'Unione ha varato, nel dicembre del 2022, il "Piano d'azione dell'UE contro il traffico di beni culturali"²¹ con l'obiettivo di rafforzare la cooperazione internazionale nella difesa del patrimonio culturale contro le attività illecite che lo minacciano, «migliorando la prevenzione e l'individuazione dei reati da parte degli operatori del mercato e delle istituzioni preposte alla tutela del patrimonio culturale»²².

¹⁸ Committee on Culture and Education, *AI in Education, Culture and the Audiovisual sector, cit.*, punto 91, https://www.europarl.europa.eu/doceo/document/TA-9-2021-0238_IT.pdf.

¹⁹ Degu Abate et al. "Artificial Intelligence to Fight Illicit Trafficking of Cultural Property," *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 48M2, (2023). DOI: 10.5194/isprs-archives-XLVIII-M-2-2023-3-2023.

²⁰ Basti pensare che, secondo quanto riportato da i Carabinieri del Comando Tutela Patrimonio Culturale, nel solo 2022 in Italia, sono stati recuperati 80.522 beni d'arte per un valore complessivo stimato di € 84.274.073. Cfr. Ministero della Cultura, "L'attività del Comando Carabinieri per la Tutela del Patrimonio Culturale nel 2022." <https://www.beniculturali.it/comunicato/24523#:~:text=Tra%20i%20recuperi%20del%202022,cui%20951%20di%20arte%20contemporanea>. Ultimo accesso 17 Gennaio 2024.

²¹ Committee on Culture and Education, *AI in Education, Culture and the Audiovisual sector, cit.*, punto 60, https://www.europarl.europa.eu/doceo/document/TA-9-2021-0238_IT.pdf.

²² Commissione Europea, "Lotta contro il traffico illecito di beni culturali." <https://culture.ec.europa.eu/it/cultural-heritage/cultural-heritage-in-eu-policies/protection-against-illicit-trafficking>. Ultimo accesso 17 Gennaio 2024.

L'intelligenza artificiale costituisce, in tale scenario, un valido alleato, se usato sapientemente, come dimostrano i casi di alcuni progetti come, ad esempio, *ENIGMA*²³ finalizzato a identificare e tracciare i beni culturali o *AURORA - Artwork Unique Recognition and Tracking*²⁴ per la combinazione di agenti chimici, tecnologie di blockchain, scansione profonda degli oggetti e sistemi miniaturizzati in grado di creare una contromisura alle truffe e agli illeciti in case d'asta, servizi logistici, commercio etc. I progetti menzionati costituiscono un esempio delle attività finanziate e avviate dall'Unione Europea per contrastare i reati connessi con il patrimonio culturale²⁵.

Infine, l'ultimo aspetto su cui le osservazioni entrano in merito, concerne l'applicazione dell'intelligenza artificiale per la creazione di contenuti artistici e culturali. A tal proposito non si può non menzionare il caso della Decima Sinfonia di Beethoven, rimasta incompiuta sino a quando, un gruppo internazionale di musicologi ed esperti di AI, sulla base degli appunti e degli scritti prodotti dall'autore della sinfonia è stato in grado di portarla a compimento nel 2021 impiegando l'intelligenza artificiale²⁶. Per favorire iniziative di tal genere, le Commissioni hanno enfatizzato la necessità di un maggior impegno verso la promozione di programmi dedicati e la creazione di «centri per la creatività digitale in cui [...] sviluppare applicazioni basate sull'IA, imparare a utilizzare queste e altre tecnologie e sperimentarne l'impiego»²⁷.

L'attenzione europea verso l'intelligenza artificiale non si esaurisce con il report bensì prosegue, specificatamente per il settore culturale, con il briefing dell'aprile 2023 in Parlamento Europeo sul tema “*Artificial intelligence in the context of cultural heritage and museums: Complex challenges and new opportunities*”²⁸. All'interno del documento di resoconto si sottolinea l'importanza dell'AI per la promozione di servizi per la salvaguardia, gestione, tutela e valorizzazione del patrimonio culturale con particolare attenzione a quello europeo, ampiamente dipendente da tecnologie e piattaforme prodotte al fuori dell'UE²⁹. Peculiari, in tal senso, sono – per metodologia e risultati – le

²³ ENIGMA. <https://eu-enigma.eu/work-packages/>. Ultimo accesso 17 Gennaio 2024.

²⁴ Aurora. <https://www.aurora-euproject.eu/>. Ultimo accesso 17 Gennaio 2024.

²⁵ Per un approfondimento sui diversi progetti finanziati dall'UE Cfr. Commissione europea, “Lotta contro il traffico illecito di beni culturali,” *cit.* <https://culture.ec.europa.eu/it/cultural-heritage/cultural-heritage-in-eu-policies/protection-against-illicit-trafficking>.

²⁶ Franco Sarcina, “L'incompiuta Decima Sinfonia di Beethoven portata a termine grazie all'Intelligenza Artificiale,” *Il Sole 24 Ore*, 30 Settembre 2021, <https://www.ilssole24ore.com/art/l-incompiuta-decima-sinfonia-beethoven-portata-termine-grazie-all-intelligenza-artificiale-AEYmrXm>.

²⁷ Committee on Culture and Education, *AI in Education, Culture and the Audiovisual sector*, *cit.*, punto 59, https://www.europarl.europa.eu/doceo/document/TA-9-2021-0238_IT.pdf.

²⁸ Parlamento Europeo, *Artificial intelligence in the context of cultural heritage and museums: Complex challenges and new opportunities* (2023),

[https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/747120/EPRS_BRI\(2023\)747120_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/747120/EPRS_BRI(2023)747120_EN.pdf)

²⁹ Come sottolineato all'interno del briefing, il 90% dei dati europei è gestito da aziende americane mentre il numero di piattaforme europee sul complessivo globale non supera il 3%. Cfr. Parlamento Europeo, *Artificial intelligence in the context of cultural heritage and museums: Complex challenges and new opportunities* (2023), p. 6. [https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/747120/EPRS_BRI\(2023\)747120_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/747120/EPRS_BRI(2023)747120_EN.pdf).

innumerevoli pratiche di digitalizzazione che implicano realtà aumentata (*Augmentity Reality – AR*)³⁰ e scansioni 3D di monumenti e artefatti, nonché tecnologie che sfruttano le reti neurali per la ricostruzione di opere d'arte³¹, testi³² o l'individuazione di nuovi siti archeologici³³.

Non sempre, tuttavia, i risultati conseguiti possono dirsi positivi e il rischio che l'utilizzo di tali tecnologie comporta, ha ripercussioni, non soltanto sul patrimonio culturale, ma sui territori e su tutto ciò che concerne la sfera dell'individuo. In particolare, è l'aspetto etico ad essere esaminato e le sfide poste in essere da un uso non appropriato dell'AI. Nel patrimonio culturale, il fattore etico è imprescindibilmente legato con l'identità e il retaggio di popoli e comunità laddove, ad esempio, l'impiego di strumenti condizionati da dati di addestramento possono risultare influenzati da correnti di pensiero³⁴ o costituire un rischio per la protezione di dati personali e sensibili.

Difatti, in un contesto in cui è necessaria una ingente quantità di dati per l'addestramento di tali sistemi, è di primaria importanza scongiurare il pericolo di una raccolta massiva di dati personali (*webscaping*) attuata con modalità poco trasparenti e senza adeguato riserbo. In Italia, lo stesso Garante per la protezione dei dati personal ha avviato, nel novembre del 2023, un'indagine conoscitiva finalizzata a verificare le misure di sicurezza adottate, su siti internet pubblici e privati, per impedire questo fenomeno. «L'indagine – si legge sul sito – coinvolge tutti i soggetti pubblici e privati, operanti quali titolari del trattamento, stabiliti in Italia o che offrono in Italia servizi, che mettono a disposizione on-line dati personali liberamente accessibili anche dagli “spider” dei produttori di algoritmi di intelligenza artificiale»³⁵. Tuttavia, non è solo nella fase di addestramento che si manifestano i pericoli connessi con la violazione dei dati personali, ma anche nella restituzione

³⁰ Lien Acke, Kristel De Vis, Stijn Verwulgen, and Jouke Verlinden, “Survey and literature study to provide insights on the application of 3D technologies in objects conservation and restoration,” *Journal of Cultural Heritage* 49, (2021). DOI: 10.1016/j.culher.2020.12.003.

³¹ È il caso del dipinto di Rembrandt *“The Night Watch”*, deturpato nel 1715 e oggetto di restauro, dal 2019, tramite reti neurali per la ricostruzione delle parti danneggiate. Attualmente, il lavoro è stato ultimato e il dipinto è tornato in esposizione al Rijksmuseum di Amsterdam. Crf. Isis Davis-Marks, “Lost Edges of Rembrandt’s ‘Night Watch’ Are Restored Using Artificial Intelligence,” *Smithsonian Mazine*, 25 Giugno 2021, <https://www.smithsonianmag.com/smart-news/lost-edges-rembrandts-night-watch-are-restored-using-artificial-intelligence-180978056/>.

³² Rodrigue Rizk, Dominick Rizk, Frederic Rizk, and Ashok Kumar, “A Hybrid Capsule Network-based Deep Learning Framework for Deciphering Ancient Scripts with Scarce Annotations: a Case Study on Phoenician Epigraphy,” *2021 IEEE International Midwest Symposium on Circuits and Systems (MWSCAS)*, (2021). DOI: 10.1109/MWSCAS47672.2021.9531798.

³³ Wouter Verschoof-van der Vaart, and Karsten Lambers, “Learning to Look at LiDAR: The Use of R-CNN in the Automated Detection of Archaeological Objects in LiDAR Data from the Netherlands,” *Journal of Computer Applications in Archeology* 2, n. 1 (2019). DOI: 10.5334/jcaa.32.

³⁴ Sofia Pensoni et al., “Artificial Intelligence and Cultural Heritage: Design and Assessment of an Ethical Framework,” *2023 IEEE International Symposium on Ethics in Engineering, Science and Technology (ETHICS)*, (2023). DOI: 10.5194/isprs-archives-XLVIII-M-2-2023-1149-2023.

³⁵ Garante per la protezione dei dati personali, “Intelligenza artificiale: Garante privacy apre un'indagine sulla raccolta di dati personali on line per addestrare gli algoritmi. L'iniziativa è volta a verificare l'adozione di misure di sicurezza da parte di siti pubblici e privati,” <https://www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/9952078>. Ultimo accesso 7 Gennaio 2023.

dei risultati con il rischio di ottenere dati personali e sensibili che non dovrebbero essere noti. Il recupero di informazioni personali e sensibili, anche oscurate, comporta un pericolo per la tutela della sfera personale di ogni individuo a cui deve essere garantito, secondo quanto stabilito all'articolo 8 della Carta dei Diritti dell'Unione Europea³⁶, la protezione dei dati di carattere personale. Le conseguenze di ciò, come sottolineato a più riprese nel citato Report del Parlamento europeo, possono avere gravi ripercussioni all'interno delle società causando disparità e discriminazioni. Ciò chiaramente non coinvolge soltanto il settore della cultura ma tutto ciò che concerne la sfera dell'individuo. In ambito sanitario, ad esempio, il rispetto della privacy è imprescindibile nell'adozione di nuove tecnologie su dati clinici di cui occorre considerare alcuni aspetti di vitale importanza come³⁷: l'esposizione di dati personali, la condivisione di dati sensibili senza il consenso informato e forme di consenso inappropriate. Il mancato o insufficiente rispetto di taluni principi ha un impatto diretto sui processi decisionali con conseguenti ripercussioni sulla salute degli individui come è stato opportunamente sottolineato dal Garante nel "Decalogo per la realizzazione di servizi sanitari nazionali attraverso sistemi di Intelligenza Artificiale"³⁸. L'esempio riportato nel Decalogo rimanda a quanto accaduto negli Stati Uniti dove il sistema di AI, utilizzato per stimare il rischio legato alla salute di oltre 200 milioni di americani, è risultato essere discriminatorio, a parità di condizioni di salute, verso i pazienti afroamericani. A seguito di una ricerca condotta sul caso³⁹, è stato difatti scoperto che il sistema tendeva ad assegnare a quest'ultimi una stima di rischio minore rispetto ad altri pazienti nelle medesime condizioni a causa delle metriche impiegate per la valutazione. Uno dei valori presi in esame era la spesa sanitaria media sostenuta, favorendo le categorie con un esborso maggiore. L'appartenenza ad una specifica etnia assume così un peso influenzando direttamente decisioni inerenti la salute degli individui. Sul piano culturale le conseguenze dettate dall'utilizzo di sistemi inficiati di bias etici e stereotipati contribuiscono ad accrescere il divario sociale tra etnie diverse favorendo in tal modo una visione distorta delle diversità culturali. Per ovviare a un tale scenario occorre considerare adeguatamente la pluralità, di lingue, etnie e culture, nei processi di implementazione delle tecnologie affinché queste possano essere uno

³⁶ Commissione Europea, 2016/C 202/389 *Charter of Fundamental Rights of the European Union* (2016), <https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:12016P/TXT>.

³⁷ Lilliana Szabo et al., "Clinician's guide to trustworthy and responsible artificial intelligence in cardiovascular imaging," *Frontiers in Cardiovascular Medicine*, (2022). DOI: 10.3389/fcvm.2022.1016032

³⁸ Garante per la protezione dei dati personali, *Decalogo per la realizzazione di servizi sanitari nazionali attraverso sistemi di Intelligenza Artificiale* (2021), <https://www.garanteprivacy.it/documents/10160/0/Decalogo+per+la+realizzazione+di+servizi+sanitari+nazionali+attraverso+sistemi+di+Intelligenza+Artificiale.pdf/a5c4a24d-4823-e014-93bf-1543f1331670?version=2.0&LinkSource=PassleApp>.

³⁹ Ziad Obermeyer et al., "Dissecting racial bias in an algorithm used to manage the health of populations," *Science* 366, (2019). DOI:10.1126/science.aax2342.

strumento di inclusione e partecipazione sociale e non costituiscono in alcun modo un pericolo per la sicurezza o la prevaricazione di popoli.

Il presente lavoro di tesi ha l'obiettivo di applicare le tecnologie intelligenza artificiale, basate sulle reti neurali, per l'estrazione della conoscenza nel dominio del patrimonio culturale e, contestualmente, proporre una modalità di disseminazione e valorizzazione delle risorse digitali e degli elementi estratti attraverso lo standard *International Image Interoperability Framework* (IIIF). La ricerca condotta durante i tre anni di dottorato è stata indirizzata, con particolare attenzione, verso l'approfondimento del framework IIIF per la descrizione standardizzata delle risorse culturali e delle *Convolutional Neural Networks* (CNNs), l'architettura di reti neurali funzionale all'analisi di immagini e audiovideo. L'approccio adottato nella ricerca mira, in prima fase, ad analizzare le tecnologie applicate per la disseminazione e valorizzazione, nel digitale, degli oggetti culturali. In tale fase la ricerca è finalizzata a individuare le modalità per la descrizione standardizzata del patrimonio culturale riconoscendo nel framework IIIF un potenziale strumento. Successivamente, l'attenzione è rivolta verso l'impiego delle citate tecnologie di reti neurali al contesto del patrimonio culturale, definendo uno stato dell'arte sulle tecniche e i casi applicativi nel dominio culturale. Il presente elaborato di tesi comprende una parte connotata da un approccio sperimentale la cui idea, consiste nell'utilizzare sistemi basati su reti neurali convoluzionali per estrarre conoscenza dalle immagini relative a oggetti del patrimonio culturale, inclusi manoscritti e documenti digitalizzati e, attraverso lo sviluppo di specifici strumenti permettere una descrizione, anche automatica, e standardizzata IIIF di tali elementi.

Per le finalità descritte, il lavoro di tesi è indirizzato in particolare verso gli istituti culturali e i *player* di mercato coinvolti nella disseminazione e comunicazione del patrimonio culturale e, in ottica più generica, interessa il territorio e gli utenti che, per motivi di studio, ricerca o lavoro possono trarre vantaggio e beneficio da tale forma di valorizzazione della conoscenza del patrimonio culturale.

Il presente lavoro di tesi si compone di cinque capitoli così organizzati:

- Capitolo 1 "Il Patrimonio culturale". Il capitolo iniziale introduce il dominio del patrimonio culturale, delineando il contesto nazionale e il ruolo dell'UNESCO come promotore, a livello internazionale, di attività di tutela e valorizzazione. Nel capitolo l'attenzione è rivolta al valore del patrimonio culturale, in particolare di quello immateriale, come fattore di aggregazione sociale e, quale elemento centrale negli asset turistici. Congiuntamente, si stabiliscono alcune considerazioni in merito alle attività di digitalizzazione, principale veicolo per la condivisione e comunicazione del patrimonio culturale, alla luce del Piano nazionale di digitalizzazione del patrimonio culturale e dei finanziamenti ad esso dedicato dal PNRR.

- Capitolo 2 “Le metodologie di descrizione standardizzata di immagini come strumento di diffusione e valorizzazione del patrimonio culturale”. Il secondo capitolo pone l’attenzione sullo standard *International Image Interoperability Framework* (IIIF) come strumento per la descrizione, disseminazione e valorizzazione del patrimonio culturale. Il framework rappresenta, a livello internazionale, una delle principali soluzioni scelte per la promozione di risorse culturali digitali, specie di natura documentale e libraria. Il capitolo ha la finalità di descrivere sommariamente le funzionalità del framework concentrando l’attenzione sulle tecnologie e strumenti che si rendono necessari per l’applicazione. In aggiunta, con l’obiettivo di evidenziare le funzionalità in continua evoluzione del IIIF, nel capitolo sono presi ad esame alcuni tra i più rappresentativi casi applicativi, nazionali ed internazionali, negli ambiti culturali MAB (Musei, Archivi e Biblioteche).
- Capitolo 3 “Analisi delle immagini e riconoscimento automatico di immagini statiche dinamiche e audiovideo: metodologie e strumenti”. Il terzo capitolo ha per oggetto le tecniche di Deep Learning (DL) per l’analisi di immagini e audiovideo. La trattazione è finalizzata a descrivere le *Convolutional Neural Networks* (CNNs) impiegate su immagini per operazioni di *Image Classification*, *Object Detection*, *Character Recognition* e sugli audiovideo per l’*Automatic Speech Recognition*. La descrizione prende in esame tali operazioni evidenziando, per ciascun di esse, alcune tra le più significative implementazioni di architetture CNNs degli ultimi anni. In particolare, nel capitolo si individuano alcuni dei principali strumenti *open source* che consentono di applicare le citate tecnologie al settore culturale, senza necessariamente disporre di elevato expertise tecnico-informatico.
- Capitolo 4 “Lo standard IIIF per la valorizzazione della conoscenza estratta automaticamente tramite Reti Neurali”. Il quarto capitolo del lavoro di tesi ha l’obiettivo di delineare una modalità di disseminazione, attraverso lo standard IIIF, della conoscenza estratta tramite le reti neurali convoluzionali. Il capitolo è caratterizzato da un approccio sperimentale circa l’applicazione delle tecnologie, tecniche e strumenti, analizzati nel corso della trattazione, al patrimonio culturale. Nel dettaglio, la prima parte del capitolo è indirizzata alle tecniche di estrazione automatica delle conoscenze, evidenziando gli aspetti necessari per l’efficace applicabilità al settore culturale. La seconda parte indica le modalità attraverso cui gli elementi estratti possono essere evidenziati, unitamente alla risorsa culturale digitale, in un modello di descrizione standardizzata secondo il IIIF.
- Capitolo 5 “Prospettive future e conclusioni”. Nel capitolo conclusivo sono pongono in essere alcune considerazioni in merito all’attuale scenario europeo, ma non solo, sui rischi connessi con l’implementazione e uso dell’intelligenza artificiale. In particolare, nel capitolo si delinea,

in maniera non troppo approfondita, il quadro politico europeo in merito alla standardizzazione dell'intelligenza artificiale individuando i principali pericoli per la sicurezza delle persone e, con particolare richiamo alla sfera culturale, per la tutela dell'identità di popoli e comunità. Infine, alla luce dei risultati evidenziati nel corso della ricerca, il capitolo termina con alcune conclusioni finali in merito al rapporto tra l'uomo e l'intelligenza artificiale, letto in un contesto interdisciplinare quali le *Digital Humanities*.

Al termine del presente lavoro di tesi sono presenti quattro Appendici come ulteriore approfondimento di taluni temi affrontati nel corso della trattazione:

- Appendice 1 “Scheda descrittiva in VRA Core de La Scuola di Atene”. L'appendice riporta, a titolo esemplificativo, la descrizione dell'opera d'arte in oggetto secondo lo standard di metadati VRA Core. Quest'ultimo, attualmente alla versione 4.0, è lo standard mantenuto dalla Library of Congress adoperato per la descrizione degli oggetti culturali visivi che permette di focalizzare l'attenzione sull'oggetto culturale sotto il duplice livello di entità e rappresentazione visiva.
- Appendice 2 “Esempio del file JSON-LD per il modello di Collection IIF”. Nella seconda appendice è riportato un esempio della rappresentazione delle risorse culturali tramite il modello IIF, secondo le Presentation API 3.0, oggetto del Capitolo 4.
- Appendice 3 “Risultati *Video Activity Recognition*”. Nell'appendice sono riportati i risultati prodotti dall'elaborazione del processo di riconoscimento delle attività tramite le reti neurali convoluzionali sui documentari realizzati dal Ministero della Cultura italiano.
- Appendice 4 “Estrazione automatica delle informazioni dai testi”. Nelle immagini e negli audiovideo, una rilevante fonte informativa è costituita dai testi che possono essere recuperati e, di conseguenza, analizzati. A tal proposito, la quarta appendice ha per oggetto la trattazione dell'estrazione automatica della conoscenza da testi descrivendo, nello specifico, le principali operazioni di *Natural Language Processing* (NLP) e alcune delle maggiori librerie e modelli oggi applicati attraverso cui analizzare efficacemente i testi per l'estrazione di entità nascoste. Infine, si descrive il processo di estrazione automatica delle informazioni da grafici attuato attraverso la combinazione di reti neurali e sistemi di riconoscimento dei caratteri. Applicate ai grafici, le tecniche di estrazione automatica delle informazioni hanno la finalità di rendere i dati *machine readable* e dunque processabili attraverso le macchine. Questa attività può risultare proficua per far emergere importanti elementi che determinano e condizionano il patrimonio culturale verso cui possono essere indirizzate attività di salvaguardia e promozione.

2. Il Patrimonio culturale

L'UNESCO - l'Organizzazione delle Nazioni Unite per l'Educazione, la Scienza e la Cultura ⁴⁰ - è l'organismo deputato, a livello internazionale, a definire linee programmatiche d'azione sul patrimonio culturale e ad assicurarne la conformità tra i Paesi. Esso ad oggi conta 193 Stati membri e 11 membri associati⁴¹. L'UNESCO definisce il patrimonio culturale come:

«Cultural heritage includes artefacts, monuments, a group of buildings and sites, museums that have a diversity of values including symbolic, historic, artistic, aesthetic, ethnological or anthropological, scientific and social significance. It includes tangible heritage (movable, immobile and underwater), intangible cultural heritage (ICH) embedded into cultural, and natural heritage artefacts, sites or monuments. The definition excludes ICH related to other cultural domains such as festivals, celebration etc. It covers industrial heritage and cave paintings»⁴².

Nel documento “2009 UNESCO Framework for Cultural Statistics (FCS)”⁴³ ritroviamo una classificazione in “domini culturali”, espressione attraverso la quale l'UNESCO intende classificare la sfera culturale distinguendo tra:

- *Cultural and Natural Heritage*: tale categoria comprende al proprio interno i beni culturali materiali come artefatti, edifici e monumenti; i beni paesaggistici quale risultato derivato dall'opera congiunta dell'uomo con la natura e come tale simbolo del legame perpetuo tra persone e territorio; e infine, i beni naturali in cui sono compresi le conformazioni geologiche e fisiche naturali e aree delimitative di habitat naturali per flora e fauna.
- *Performance and Celebration*: sono incluse tutte le forme di espressioni di eventi culturali dal vivo includendo concerti, teatri, festival e così via.
- *Visual Arts and Crafts*: sono riconducibili a tale categoria la creazione di opere che hanno una manifestazione visiva reale come dipinti, sculture, oggetti di artigianato etc. elaborati per mezzo di un lavoro manuale.
- *Books and Press*: include ogni forma di pubblicazione, analogica o informatica, di libri, giornali o periodici.

⁴⁰ United Nations for Education, Science and Culture (UNESCO), “UNESCO in brief,” ultimo accesso 23 Gennaio 2024, <https://unesco.org/en/brief>.

⁴¹ Consiglio d'Europa, “Organizzazione delle Nazioni Unite per l'Educazione, la Scienza e la Cultura (UNESCO),” ultimo accesso 23 Gennaio 2024, <https://coe.int/it/web/cultural-routes/unesco>.

⁴² United Nations for Education, Science and Culture (UNESCO) Glossary, “Cultural Heritage,” ultimo accesso 23 Gennaio 2024, https://uis.unesco.org/en/glossary/cultural_heritage..

⁴³ United Nations for Education, Science and Culture Institute for Statistics (UNESCO UIS), and UNESCO Culture Sector, *2009 UNESCO Framework for Cultural Statistics (FCS)* (2009), https://uis.unesco.org/sites/default/files/documents/unesco-framework-for-cultural-statistics-2009-en_0.pdf. Questo documento costituisce una versione revisionata di quello edito nel 1986.

- *Audio-visual and Interactive Media*: tale categoria include le opere e gli oggetti elettronici non riconducibili alla categoria di Visual Arts and Crafts come: film, videogiochi, forme di espressione culturale via web (social network, blog, siti etc.) e prodotti che possono essere ricondotti alla categoria di “arte virtuale”.
- *Design and Creative Services*: sono ricomprese le attività, oggetti e servizi risultati da un processo di realizzazione creativo, artistico ed estetico come, ad esempio, la realizzazione di spot pubblicitari o il disegno di un progetto architettonico.

La distinzione degli ambiti culturali operata in tale sede, finalizzata a valutare le attività e i servizi culturali offerti in ciascun settore, trova il proprio fondamento sulla nozione di cultura riaffermata nella Dichiarazione Universale sulla Diversità Culturale⁴⁴:

«Culture should be regarded as the set of distinctive spiritual, material, intellectual and emotional features of society or a social group, and that it encompasses, in addition to art and literature, lifestyles, ways of living together, value systems, traditions and beliefs»⁴⁵.

Sono dunque i valori incarnati dalla cultura che, veicolati attraverso luoghi, oggetti fisici e forme astratte, contribuiscono a dare una identità a popoli e comunità, in modo che gli individui possano sviluppare un comune senso di appartenenza. In tal senso l'identità è allora parte essenziale del concetto di diversità culturale, intesa come un insieme di *«uniqueness and plurality of the identities of the groups and societies making up humankind»⁴⁶*. Sul concetto di diversità culturale, l'UNESCO pone un duplice accento: da un lato come strumento, per mezzo del quale è possibile riconoscere e legittimare l'esistenza di altre comunità e, dall'altro, come diritto, ossia l'imperativo etico inseparabile dal rispetto della dignità umana⁴⁷. In tale cornice il patrimonio culturale costituisce, dunque, il medium attraverso cui la diversità si manifesta e le comunità si riconoscono.

Attorno l'idea di cultura, l'UNESCO ha proceduto nel corso della sua attività a individuare le peculiarità che contraddistinguono il patrimonio culturale e si rendono necessarie per la sua salvaguardia. Nella “Convenzione riguardante la protezione del patrimonio culturale e naturale”⁴⁸ del 1972, l'UNESCO ha riconosciuto due principali categorie di patrimonio culturale:

⁴⁴ United Nations for Education, Science and Culture (UNESCO), *UNESCO Universal Declaration on Cultural Diversity* (2001), <https://en.unesco.org/about-us/legal-affairs/unesco-universal-declaration-cultural-diversity>.

⁴⁵ United Nations for Education, Science and Culture Institute for Statistics (UNESCO UIS), and UNESCO Culture Sector, *2009 UNESCO Framework for Cultural Statistics (FCS) cit.*, p. 19, https://uis.unesco.org/sites/default/files/documents/unesco-framework-for-cultural-statistics-2009-en_0.pdf

⁴⁶ United Nations for Education, Science and Culture (UNESCO), *UNESCO Universal Declaration on Cultural Diversity cit.*, art. 1, <https://en.unesco.org/about-us/legal-affairs/unesco-universal-declaration-cultural-diversity>.

⁴⁷ United Nations for Education, Science and Culture (UNESCO), *UNESCO Universal Declaration on Cultural Diversity cit.*, artt. 3-4, <https://en.unesco.org/about-us/legal-affairs/unesco-universal-declaration-cultural-diversity>.

⁴⁸ United Nations for Education, Science and Culture (UNESCO), *Convention Concerning the Protection of the World Cultural and Natural Heritage* (1972), <https://whc.unesco.org/en/conventiontext/>.

- patrimonio culturale;
- patrimonio naturale.

I primi sono individuati all'art.1 della Convenzione e distinti in:

- Monumenti: rientrano in tale categoria le «opere architettoniche, plastiche o pittoriche monumentali, elementi o strutture di carattere archeologico, iscrizione e grotte e gruppi di elementi dal valore universale eccezionale dall'aspetto storico, artistico o scientifico»;
- Agglomerati: con tale accezione si prende a riferimento «gruppi di costruzioni isolate o riunite che, per la loro architettura, unità o integrazione nel paesaggio hanno un valore universale eccezionale dall'aspetto storico, artistico o scientifico»;
- Siti: in tale categoria rientrano «le opere dell'uomo o coniugate dell'uomo e della natura, come anche le zone, compresi i siti archeologici, di valore universale eccezionale dall'aspetto storico ed estetico, etnologico o antropologico»⁴⁹.

I beni del patrimonio naturale sono invece oggetto dell'art. 2 della Convenzione e classificati nelle medesime categorie dei corrispettivi culturali, a differenza dei quali, emerge la specificità del fattore naturale:

- Monumenti naturali: «costituiti da formazioni fisiche e biologiche o da gruppi di tali formazioni di valore universale eccezionale dall'aspetto estetico o scientifico»;
- Formazioni geologiche e fisiche: insieme alle «zone strettamente delimitate costituenti l'habitat di specie di animali e vegetali minacciate, di valore universale eccezionale dall'aspetto scientifico o conservativo»;
- Siti naturali o zone naturali: anch'esse «strettamente delimitate di valore universale eccezionale dall'aspetto scientifico, conservativo o estetico naturale»⁵⁰.

La distinzione operata nella Convenzione risulta essere, tuttavia, riduttiva del concetto stesso di *patrimonio culturale* del quale sono riconosciuti unicamente beni di tipo materiale escludendo, invece, la componente intangibile. A partire dagli anni successivi ha inizio un processo di valorizzazione degli aspetti immateriali della cultura, distinti dai beni materiali e naturali, quale principale espressione della diversità culturale. Tale processo conduce alla redazione, nel 1989, delle "Raccomandazioni sulla salvaguardia della cultura tradizionale e del folklore"⁵¹ nelle quali si

⁴⁹ United Nations for Education, Science and Culture (UNESCO), *Convention Concerning the Protection of the World Cultural and Natural Heritage cit.*, art. 1, <https://whc.unesco.org/en/conventiontext/>.

⁵⁰ United Nations for Education, Science and Culture (UNESCO), *Convention Concerning the Protection of the World Cultural and Natural Heritage cit.*, art. 2, <https://whc.unesco.org/en/conventiontext/>.

⁵¹ United Nations for Education, Science and Culture (UNESCO), *Recommendation on the Safeguarding of Traditional Culture and Folklore* (1989), <https://www.unesco.org/en/legal-affairs/recommendation-safeguarding-traditional-culture-and-folklore>.

attribuisce alla cultura popolare e tradizionale la funzione di espressione culturale e, come tale, soggetta ai medesimi interessi del patrimonio culturale e naturale. L'adozione, il 17 ottobre 2003, della "Convenzione per la salvaguardia del patrimonio culturale immateriale"⁵² costituisce uno spartiacque nell'ambito della cultura tramite il pieno riconoscimento del patrimonio culturale immateriale distinto dal patrimonio culturale materiale.

Rimanendo invariate le definizioni sul patrimonio naturale e materiale, la Convenzione sul patrimonio immateriale indica quest'ultimo come:

«le prassi, le rappresentazioni, le espressioni, le conoscenze, il know-how – come pure gli strumenti, gli oggetti, i manufatti e gli spazi culturali associati agli stessi – che le comunità, i gruppi e in alcuni casi gli individui riconoscono in quanto parte del loro patrimonio culturale»⁵³.

Il patrimonio immateriale comprende dunque:

- tradizioni ed espressioni orali, ivi compreso il linguaggio, in quanto veicolo del patrimonio culturale immateriale;
- le arti dello spettacolo;
- le consuetudini sociali, gli eventi rituali e festivi;
- le cognizioni e le prassi relative alla natura e all'universo;
- l'artigianato tradizionale.

Il patrimonio culturale, inteso nella sua totalità materiale e immateriale, tiene traccia del percorso evolutivo dell'uomo e costituisce memoria anche delle relazioni con il territorio e l'ambiente circostante.

L'UNESCO mette in risalto il ruolo della cultura quale strumento di coesione sociale e, come tale, invoca la collaborazione internazionale di Stati e Governi ad attuare politiche in favore del patrimonio culturale. Ciascun Stato membro dell'UNESCO, aderendo alle Convenzioni, riconosce quanto esplicitamente manifestato attuando misure nazionali nel rispetto di ciò. In ambito nazionale, il patrimonio culturale è regolamentato attraverso il Codice dei beni culturali e del paesaggio⁵⁴ in cui si stabilisce, all'art. 2 comma 1, che:

⁵² United Nations for Education, Science and Culture (UNESCO), *Convention for the Safeguarding of the Intangible Cultural Heritage* (2003), <https://ich.unesco.org/en/convention>.

⁵³ United Nations for Education, Science and Culture (UNESCO), *Convenzione per la salvaguardia del patrimonio culturale immateriale*, cit., tradotto da Commissione Nazionale Italiana per UNESCO, art. 2, <https://ich.unesco.org/doc/src/00009-IT-PDF.pdf>.

⁵⁴ D. Lgs. 42/2004 *Codice dei beni culturali e del paesaggio* (2004), <https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:decreto.legislativo:2004-01-22;42>.

«Il patrimonio culturale è costituito dai beni culturali e paesaggistici» intendendo per beni culturali «le cose mobili e immobili che [...] presentano interesse artistico, storico, archeologico, etnoantropologico, archivistico e bibliografico e le altre cose individuate dalla legge quali testimonianza di valore di civiltà»⁵⁵.

Pur nel rispetto delle Convenzioni UNESCO, il Codice non attua una netta distinzione tra patrimonio materiale e immateriale riconducendoli entrambi sotto la dicitura di *beni culturali* stabilendo, all'art. 7 bis che:

«Le espressioni di identità culturale collettiva contemplate dalle Convenzioni UNESCO per la salvaguardia del patrimonio culturale immateriale e per la protezione e la promozione delle diversità culturali [...] sono assoggettabili alle disposizioni del presente codice qualora siano rappresentate da testimonianze materiali e sussistano i presupposti e le condizioni per l'applicabilità dell'articolo 10»⁵⁶.

I beni culturali interessati dalle disposizioni dettate dal Codice all'art.10 sono, dunque, l'insieme delle cose mobili e immobili di dominio pubblico⁵⁷ che manifestano interesse storico, artistico, archeologico o etnoantropologico, come:

- Raccolte di musei, pinacoteche, gallerie ed altri luoghi espositivi;
- Archivi e singoli documenti appartenenti allo Stato o altri soggetti pubblici;
- Raccolte librerie⁵⁸.

Ulteriormente, il Codice disciplina anche i beni appartenenti a privati che, ove sia espressa la “dichiarazione in interesse culturale”⁵⁹, si configurano come beni culturali e, come tali, disciplinati dal Codice⁶⁰.

Rispetto a quanto svolto dall'UNESCO sul piano classificatorio, in Italia si è proceduto verso una più scrupolosa riorganizzazione del patrimonio naturale e culturale, materiale e immateriale. L'Istituto Centrale per il Catalogo e la Documentazione (ICCD), uno degli Istituti del Ministero

⁵⁵ D. Lgs. 42/2004 *Codice dei beni culturali e del paesaggio cit.*, art. 2 comma 1, <https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:decreto.legislativo:2004-01-22;42>.

⁵⁶ D. Lgs. 42/2004 *Codice dei beni culturali e del paesaggio cit.*, art. 7 bis <https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:decreto.legislativo:2004-01-22;42>.

⁵⁷ Con tale accezione si rimanda a quanto puntualmente stabilito dal D. Lgs. 42/2004 *Codice dei beni culturali e del paesaggio cit.*, art. 10 comma 1: «appartenenti allo Stato, alle regioni, agli altri enti pubblici territoriali, nonché ad ogni altro ente ed istituto pubblico e a persone giuridiche private senza fine di lucro, ivi compresi gli enti ecclesiastici civilmente riconosciuti».

⁵⁸ D. Lgs. 42/2004 *Codice dei beni culturali e del paesaggio cit.*, art. 10 comma 2, <https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:decreto.legislativo:2004-01-22;42>.

⁵⁹ D. Lgs. 42/2004 *Codice dei beni culturali e del paesaggio cit.*, art. 13, <https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:decreto.legislativo:2004-01-22;42>.

⁶⁰ D. Lgs. 42/2004 *Codice dei beni culturali e del paesaggio cit.*, art. 10 comma 3, <https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:decreto.legislativo:2004-01-22;42>.

della Cultura (MiC) dotati di autonomia speciale⁶¹ per la promozione e valorizzazione del patrimonio culturale attraverso la catalogazione e l'elaborazione di normative in materia, ha classificato i beni culturali in base alla tipologia e sfera di appartenenza individuando un totale di nove “settori disciplinari” così distinti:

- Beni archeologici;
- Beni architettonici e paesaggistici;
- Beni demoetnoantropologici;
- Beni fotografici;
- Beni musicali;
- Beni naturalistici;
- Beni numismatici;
- Beni scientifici e tecnologici;
- Beni storico artistici.

Inoltre, a seconda della specificità di ciascuna risorsa, i beni possono essere a loro volta raggruppati in: beni mobili, beni immobili e beni immateriali. Per ciascuna tipologia di bene è stata creata una specifica scheda di catalogo volta a garantire una precisa e puntuale descrizione di ciascuno di essi. Allo stato attuale⁶² sono trenta le schede di catalogo create suddivise in:

tipo scheda	definizione	settore disciplinare	versione
A	<i>Architettura</i>	beni architettonici e paesaggistici	3.0
AT	<i>Reperti antropologici</i>	beni archeologici	3.01
BDI	<i>Beni demoetnoantropologici immateriali</i>	beni demoetnoantropologici	4.0
BDM	<i>Beni demoetnoantropologici materiali</i>	beni demoetnoantropologici	4.0
BNB	<i>Beni naturalistici-Botanica</i>	beni naturalistici	3.01
BNM	<i>Beni naturalistici-Mineralogia</i>	beni naturalistici	3.01
BNP	<i>Beni naturalistici-Paleontologia</i>	beni naturalistici	3.01
BNPE	<i>Beni naturalistici-Petrologia</i>	beni naturalistici	3.01

⁶¹ Attualmente afferisce all'Istituto per la Digitalizzazione del Patrimonio culturale – Digital Library. Cfr. Ministero della Cultura (MiC), “Organigramma,” ultimo accesso 23 Gennaio 2024, https://media.beniculturali.it/mibac/files/boards/be78e33bc8ca0c99bff70aa174035096/PDF/2022/organigramma2021_navigabile.pdf.

⁶² Istituto Centrale per il Catalogo e la Documentazione, “Schede di catalogo,” ultimo accesso 23 Gennaio 2024, <https://github.com/ICCD-MiBACT/Standard-catalografici/tree/master/schede-di-catalogo>.

BNPL	<i>Beni naturalistici-Planetologia</i>	beni naturalistici	3.01
BNZ	<i>Beni naturalistici-Zoologia</i>	beni naturalistici	3.01
CA	<i>Complessi archeologici</i>	beni archeologici	3.0
CNS ^[^1]	<i>Centri/nuclei storici</i>	beni architettonici e paesaggistici	N.D.
D	<i>Disegni</i>	beni storici e artistici	3.0
F	<i>Fotografia</i>	beni fotografici	4.0
FF	<i>Fondi fotografici</i>	beni fotografici	4.0
MA	<i>Monumenti archeologici</i>	beni archeologici	3.0
MI	<i>Matrici incise</i>	beni storici e artistici	3.0
NU	<i>Beni numismatici</i>	beni numismatici	3.0
OA	<i>Opere/oggetti d'arte</i>	beni storici e artistici	3.0
OAC	<i>Opere/oggetti d'arte contemporanea</i>	beni storici e artistici	3.0
PG	<i>Parchi/giardini</i>	beni architettonici e paesaggistici	3.0
PST	<i>Patrimonio scientifico e tecnologico</i>	beni scientifici e tecnologici	4.0
RA	<i>Reperti archeologici</i>	beni archeologici	3.0
S	<i>Stampe</i>	beni storici e artistici	3.0
SAS	<i>Saggi stratigrafici</i>	beni archeologici	3.0
SI	<i>Siti archeologici</i>	beni archeologici	3.0
SM	<i>Strumenti musicali</i>	beni musicali	4.0
SMO	<i>Strumenti musicali-Organo</i>	beni musicali	3.01
TMA	<i>Tabella materiali archeologici</i>	beni archeologici	3.0
VeAC	<i>Vestimenti antichi/contemporanei</i>	beni storici e artistici	3.01

Tabella 1: schede di catalogo ICCD⁶³.

Le schede di catalogo costituiscono il principale punto di riferimento nazionale, insieme a thesauri e liste controllate di termini, per la descrizione e catalogazione dei beni culturali e dalle quali attinge informazioni il Catalogo Generale dei Beni Culturali. Il Catalogo è lo strumento realizzato

⁶³ Istituto Centrale per il Catalogo e la Documentazione, "Schede di catalogo," ultimo accesso 23 Gennaio 2024, <https://github.com/ICCD-MiBACT/Standard-catalografici/tree/master/schede-di-catalogo>.

dall'ICCD per la condivisione dei dati aperti sul patrimonio culturale italiano presenti in rete. Si basa sull'impiego di *linked open data* (LOD), la «tecnologia e l'insieme di buone pratiche»⁶⁴ su cui si fonda il “web semantico” per la pubblicazione e il collegamento di dati aperti, leggibili e interpretabili dalle macchine (*machine readable*). Al fine di soddisfare i requisiti imposti dal web semantico, i dati relativi ai beni culturali confluiti nel Catalogo sono posti in connessione con altre banche-dati della rete italiana dei beni culturali⁶⁵. In tale ambito, i LOD costituiscono uno valido strumento che, adeguatamente applicato, permette di arricchire e mantenere costantemente aggiornata la base di conoscenza attorno al patrimonio culturale. Inoltre, il collegamento tra le banche dati e la circolazione di dati aperti agevolano la condivisione di informazioni tra i diversi *stakeholders* del patrimonio culturale permettendo la messa in opera di progetti finalizzati alla tutela e valorizzazione di quest'ultimo. Sono queste attività essenziali nell'ambito del patrimonio culturale, e per le quali lo Stato assume un ruolo centrale finalizzato a garantire l'assoluta promozione della cultura. Tale precetto è decretato all'interno dello stesso Codice dei beni culturali e del paesaggio, nel quale è fatto esplicito richiamo⁶⁶ all'art.9 della Costituzione in cui si stabilisce che «La Repubblica promuove lo sviluppo della cultura e la ricerca scientifica e tecnica. Tutela il paesaggio e il patrimonio storico e artistico della Nazione»⁶⁷.

La tutela deve essere, dunque, esercitata al fine di individuare le forme del patrimonio culturale e attuare le azioni necessarie a proteggerlo e preservarlo. Come tale, è un'attività estremamente complessa che vede l'implicazione di fattori naturali e prevedibili, come l'usura o il degrado dovuto al trascorrere del tempo, insieme a fattori improvvisi e violenti come ad esempio le calamità naturali, scoppio di guerre o determinati attacchi volti a deturpare o distruggere i beni culturali quale simbolo di ideologie o civiltà avverse.

La valorizzazione è l'attività, complementare alla tutela, finalizzata a promuovere il patrimonio culturale affinché sia reso conoscibile e accessibile da un maggior numero possibile di individui. Per tale ragione, la valorizzazione non può realizzarsi senza un'adeguata attività di tutela volta a garantire l'esistenza di un bene da promuovere e, parimenti, richiede una stretta collaborazione tra gli attori interessati dalla cultura. Tali attività richiedono, improrogabilmente, la collaborazione tra soggetti privati e Stato, come auspicato dall'UNESCO in ottica internazionale, finalizzata a

⁶⁴ Mauro Guerrini, and Tiziana Possemato, *Linked Data per biblioteche, archivi e musei* (Milano: Editrice Bibliografica, 2015), p. 28. ISBN: 978-8870758306.

⁶⁵ Ministero della Cultura, “Dati della cultura,” ultimo accesso 23 Gennaio 2024, <https://www.beniculturali.it/dati-della-cultura>.

⁶⁶ D. Lgs. 42/2004 Codice dei beni culturali e del paesaggio cit., art. 1 cc. 1-2, <https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:decreto.legislativo:2004-01-22;42>.

⁶⁷ Costituzione della Repubblica Italiana, art. 9, <https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:costituzione:1947-12-27~art9>.

rendere il patrimonio culturale, quale bene appartenente alla comunità, liberamente fruibile. La cooperazione è quantomai necessaria considerando l'importanza che la cultura riveste sotto l'aspetto umanistico, tanto quanto socio-economico, e per la crescita dei territori attraverso il turismo e l'industria creativa culturale. I valori incarnati dal patrimonio culturale e l'importanza che questi rivestono a fondamento di talune comunità, rappresentano il principio cardine alla base delle attività di tutela e valorizzazione che rendono necessaria la cooperazione di tutti gli attori coinvolti nell'ecosistema culturale.

L'UNESCO, nella sua missione di promozione della cultura, ha da sempre rivestito il ruolo di guida a livello internazionale avviando attività, di intesa con gli Stati membri, finalizzate a proteggere, tutelare e valorizzare il patrimonio culturale. Ciò si evidenzia nelle diverse Convenzioni emanate, specificatamente, per il patrimonio culturale e naturale mondiale⁶⁸, le diversità culturali⁶⁹, il patrimonio immateriale⁷⁰ e il patrimonio culturale subacqueo⁷¹. Le Convenzioni si configurano come lo strumento attraverso le quali gli Stati membri riconoscono il proprio ruolo di "garanti" nei confronti del relativo patrimonio culturale, adottando le misure necessarie a tutelarlo e valorizzarlo. Per evidenziare la necessità della collaborazione internazionale, con la Convenzione sulla Protezione del Patrimonio Mondiale, culturale e naturale dell'Umanità, l'UNESCO ha istituito la Lista del Patrimonio Culturale Mondiale (*World Heritage List*)⁷² in cui vi sono registrati i siti culturali proposti dai singoli Stati membri. La Lista identifica quei beni che, considerato il loro altissimo valore culturale, assumono rilevanza per l'intera umanità. Inoltre, essa permette di identificare quei beni posti in situazione di pericolo e, dunque, necessitano di maggior tutela. Similarmente a quanto avvenuto con il patrimonio culturale e naturale, con la Convenzione per la salvaguardia del Patrimonio Culturale Immateriale è stata definita la Lista del Patrimonio Culturale Immateriale (*List of Intangible Cultural Heritage*)⁷³ allo scopo di dare rilievo alle usanze, *know-how*, tradizioni che per loro caratteristica identificano le comunità. Si tratta di una categoria di patrimonio culturale per cui è necessario adottare misure in grado di tenere traccia e memoria, specialmente per ciò che concerne comunità più piccole e

⁶⁸ United Nations for Education, Science and Culture (UNESCO), *Convention Concerning the Protection of the World Cultural and Natural Heritage cit.*, <https://whc.unesco.org/en/conventiontext/>.

⁶⁹ United Nations for Education, Science and Culture (UNESCO), *UNESCO Universal Declaration on Cultural Diversity cit.*, <https://en.unesco.org/about-us/legal-affairs/unesco-universal-declaration-cultural-diversity>.

⁷⁰ United Nations for Education, Science and Culture (UNESCO), *Convention for the Safeguarding of the Intangible Cultural Heritage cit.*, <https://ich.unesco.org/en/convention>.

⁷¹ United Nations for Education, Science and Culture (UNESCO), *Convention on the Protection of the Underwater Cultural Heritage (2001)*, <https://unesdoc.unesco.org/ark:/48223/pf0000126065>.

⁷² United Nations for Education, Science and Culture (UNESCO), "World Heritage List," ultimo accesso 17 Gennaio 2024, <https://whc.unesco.org/>.

⁷³ United Nations for Education, Science and Culture (UNESCO), "List of Intangible Cultural Heritage," ultimo accesso 17 Gennaio 2024, <https://ich.unesco.org/en/lists>.

isolate. L'iscrizione nella Lista consente, a tal fine, di individuare e documentare saperi e conoscenze che, altrimenti, rischierebbero di andare persi per sempre.

Patrimonio culturale, identità dei popoli e attrattività turistica dei territori

Sul piano culturale, le liste permettono di conoscere beni, luoghi e tradizioni che caratterizzano e legano le popolazioni tra di loro e, al contempo, consentono ai Governi di attuare azioni volte a tutelare quei beni a rischio di estinzione. Uno dei vantaggi consequenziali delle liste è, senza dubbio, quello di promuovere a livello internazionale i beni iscritti caratterizzando, di conseguenza, il turismo verso i Paesi di pertinenza. L'Italia è sicuramente tra le Nazioni che maggiormente possono beneficiare del patrimonio culturale, con il maggior numero di beni censiti nella Lista del Patrimonio Culturale Mondiale, ben 59 siti culturali⁷⁴ distinti tra beni culturali (53) e beni naturali (6), insieme a 19 elementi della Lista del Patrimonio Culturale Immateriale⁷⁵. Oltre a quelli già individuati e universalmente riconosciuti, tuttavia, esistono innumerevoli beni sommersi, o scarsamente conosciuti, posti a diversi fattori di rischio per i quali l'estrazione automatica della conoscenza costituisce una modalità attraverso cui essere degnamente valorizzati. Ciò assume una maggiore rilevanza considerando l'importanza strategica che i beni culturali e naturali costituiscono per il settore del turismo, del quale rappresentano le maggiori fonti di attrazione. I dati derivanti dall'indagine sul turismo in Italia durante il 2022, realizzata da ISNART (Istituto Nazionale Ricerche Turistiche) e dall'Osservatorio sull'Economia del Turismo e delle Camere di Commercio, conferma il potere attrattivo dei beni culturali nei flussi turistici⁷⁶. Per un Paese come l'Italia, in cui il turismo costituisce uno dei principali asset economici, le tecniche di estrazione automatica della conoscenza possono essere veicolate verso la (ri)scoperta, la promozione e valorizzazione del proprio patrimonio culturale con un impatto positivo sotto i diversi aspetti umanistici, sociali ed economici.

In tale scenario, l'evoluzione tecnologica ha radicalmente mutato il paradigma socio-culturale, creando una dimensione digitale che rivoluziona le modalità attraverso cui entrare in contatto con il patrimonio. Nella sfera digitale gli strumenti di ICT (*Information and Communication Technologies*) assumono un ruolo di rilievo dalle potenzialità infinite. Rimane, tuttavia, in capo agli individui la responsabilità circa le modalità e finalità cui indirizzare tali strumenti. È

⁷⁴ United Nations for Education, Science and Culture (UNESCO), "World Heritage List," ultimo accesso 17 Gennaio 2024, https://whc.unesco.org/en/list/?search=&id_states=it&components=0&order=country.

⁷⁵ United Nations for Education, Science and Culture (UNESCO), "Intangible Heritage List," ultimo accesso 17 Gennaio 2024, <https://ich.unesco.org/en/state/italy-IT?info=elements-on-the-lists>.

⁷⁶ Istituto Nazionale Ricerche Turistiche (ISNART), "Report indagine sul turismo nazionale 2022," <https://www.isnart.it/it/economia-del-turismo/indagini-nazionali/indagine-2022/>. Ultimo accesso 17 gennaio 2024.

innegabile, di certo, attestare i vantaggi che le tecnologie di ICT permettono di raggiungere nei diversi campi di applicazione; basti pensare, ad esempio, alle avanzate tecniche di analisi che consentono di intervenire preventivamente per contrastare il deterioramento dei reperti, o alla realizzazione di apparecchiature per la realtà aumentata (*Augmentaty Reality – AR*) e virtuale (*Virtual Reality – VR*) in grado di offrire un maggior livello di attrazione e di *engagement*, soprattutto nei confronti delle nuove generazioni⁷⁷. In particolare, fra le innovazioni derivanti dallo sviluppo tecnologico, le tecniche per l'estrazione automatica della conoscenza hanno le potenzialità per costituire uno strumento basilare nei progetti e negli asset legati al patrimonio culturale. L'applicazione di tecniche e tecnologie in grado di automatizzare i processi di estrazione delle informazioni costituisce un primo step per la condivisione e disseminazione della cultura che, mediante gli strumenti del Web e della comunicazione, può essere veicolata verso i vari strati della società.

Il processo di digitalizzazione del patrimonio culturale in Italia.

Per i Governi di tutto il mondo è essenziale riconoscere le opportunità e le conseguenze legate all'uso delle tecnologie sul patrimonio culturale, per essere in grado di attirare investimenti e promuovere attività destinate a consolidarne l'uso. Nel contesto italiano, in considerazione dell'impatto che le tecnologie sull'intero ambito culturale, è stato destinato oltre 1 miliardo di Euro verso la cultura tramite la Missione 1 (Digitalizzazione, innovazione, competitività, cultura e turismo), Componente 3 (Turismo e Cultura 4.0)⁷⁸. Tra gli obiettivi del Piano MIC3 rientra la nascita di nuovi servizi e la promozione di attività per favorire partecipazione sociale al patrimonio, anche, attraverso modalità digitali. In particolare, è il sub investimento 1.1 “Piattaforme e strategie digitali per l'accesso al patrimonio culturale” ad essere destinato alla digitalizzazione del patrimonio culturale italiano ed entro il quale è stato redatto, dall'Istituto Centrale per la Digitalizzazione del Patrimonio Culturale – Digital Library, il Piano nazionale di digitalizzazione del patrimonio culturale (PND)⁷⁹.

A partire da agosto 2022 sono stati avviati progetti di interesse nazionale in collaborazione con Regioni ed Enti finalizzati a digitalizzare le collezioni di archivi, musei e biblioteche. Allo stato attuale⁸⁰, sono quattro i principali progetti portati avanti nel contesto dell'investimento 1.1.5 del

⁷⁷ Gabriel Boboc Răzvan et al., "Augmented Reality in Cultural Heritage: An Overview of the Last Decade of Applications," *Applied Sciences* 12, n. 19: 9859 (2022). DOI: 10.3390/app12199859.

⁷⁸ Ministero della Cultura, "Cultura 4.0," <https://pnrr.cultura.gov.it/>. Ultimo accesso 17 Gennaio 2024.

⁷⁹ Istituto Centrale per la Digitalizzazione del Patrimonio Culturale – Digital Library "Il Piano," <https://digitalibrary.cultura.gov.it/il-piano/>. Ultimo accesso 17 Gennaio 2024.

⁸⁰ Ministero della Cultura, "Sub-investimento 1.1.5 – Digitalizzazione," ultimo accesso 27 Gennaio 2024, <https://pnrr.cultura.gov.it/misura-1-patrimonio-culturale-per-la-prossima-generazione/1-1-piattaforme-e-strategie-digitali-per-laccesso-al-patrimonio-culturale/>.

PNRR Cultura con l'obiettivo di rendere accessibile, insieme ai propri metadati descrittivi, le risorse culturali custodite sul territorio nazionale all'interno della Digital Library italiana:

- “Microfilm”⁸¹: il progetto è finalizzato a digitalizzare i 107 mila microfilm contenenti decine di migliaia di manoscritti custoditi in 180 biblioteche italiane e 16 straniere. Uno dei principali vantaggi derivanti dalla digitalizzazione dei microfilm consiste nel tutelare le bobine attualmente conservate presso il Centro Nazionale per lo Studio del Manoscritto (CNSM). Queste rappresentano testimonianze uniche di manoscritti attualmente non più esistenti o, in taluni casi, irreparabilmente danneggiati da non poter essere movimentati o consultabili.
- “Carta: documentazione catastale (registri e mappe) e giornali quotidiani postunitari”⁸²: il progetto implica due attività distinte volte alla digitalizzazione del patrimonio documentario e giornalistico custodito negli Archivi di Stato e Biblioteche Nazionali italiane. La prima parte del progetto ha l'obiettivo la digitalizzazione di milioni di pagine della rassegna stampa post unitaria quali: il primo numero del Corriere della Sera, le cronache rosa degli Anni Cinquanta, i risultati sportivi dei giornali locali etc. La rassegna stampa oggetto del progetto copre il primo secolo di storia dell'Italia unitaria (1861-1955) attualmente conservata presso le Biblioteche Nazionali Centrali di Roma e Firenze e le Biblioteche Nazionali di Milano e di Napoli. La seconda parte del progetto mira invece alla digitalizzazione degli archivi catastali di 19 capoluoghi di Regione di tutta Italia. Il progetto prevede la realizzazione di 21 milioni di nuove digitalizzazioni di documenti d'archivio quali registri, inventari, carte sciolte e mappe catastali che testimoniano la cultura, la società e la storia del Paese.
- “Archivi fotografici: positivi, negativi, unicum e disegni”⁸³: il progetto è dedicato alla digitalizzazione di oltre 5,5 milioni di fotografie e circa 500.000 disegni degli archivi fotografici delle Soprintendenze Archeologia, Belle Arti e Paesaggio (ABAP). Si tratta di un imponente intervento volto «all'acquisizione massiva, per mezzo di scanner e macchine fotografiche, di stampe fotografiche, diapositive, negativi su pellicola e su lastra di vetro, e

⁸¹ Istituto Centrale per la Digitalizzazione del Patrimonio Culturale – Digital Library, “MICROFILM: al via la digitalizzazione dei manoscritti,” ultimo accesso 27 Gennaio 2024, <https://pnrr.cultura.gov.it/microfilm-al-via-la-digitalizzazione-dei-manoscritti-m1c31-1-strategie-e-piattaforme-digitali-per-il-patrimonio-culturale/>.

⁸² Istituto Centrale per la Digitalizzazione del Patrimonio Culturale – Digital Library, “PNRR, digitalizzazione del patrimonio culturale: online la procedura di gara da 27,9 milioni di euro per la categoria ‘Carta: documentazione catastale (registri e mappe) e giornali quotidiani postunitari’,” ultimo accesso 27 Gennaio 2024, <https://digitalibrary.cultura.gov.it/avvisi/pnrr-digitalizzazione-del-patrimonio-culturale-online-la-procedura-di-gara-da-279-milioni-di-euro-per-la-categoria-carta-documentazione-catastale-registri-e-mappe-e-giornali-quotidiani/>.

⁸³ Istituto Centrale per la Digitalizzazione del Patrimonio Culturale – Digital Library, “AVVISI | NOTIZIE | PNRR PNRR, digitalizzazione del patrimonio culturale: online la procedura di gara da 16,9 milioni di euro per la categoria ‘Archivi fotografici: positivi, negativi, unicum e disegni’,” ultimo accesso 27 Gennaio 2024, <https://digitalibrary.cultura.gov.it/notizie/pnrr-digitalizzazione-del-patrimonio-culturale-online-la-procedura-di-gara-da-169-milioni-di-euro-per-la-categoria-archivi-fotografici-positivi-negativi-unicum-e-disegni/>.

pezzi unici, anche antichi»⁸⁴. In aggiunta al materiale proveniente dalle Soprintendenze ABAP, verranno digitalizzati anche altri fondi quali, ad esempio, quelli custoditi presso l'Aero Fototeca Nazionale, l'ICCD (Istituto Centrale per il Catalogo e la Documentazione) e alcune sedi museali.

- “Oggetti museali: beni di deposito (storico-artistici, archeologici) e grafici”⁸⁵: il quarto progetto avviato dalla Digital Library del Ministero della Cultura, nell'ambito delle attività previste dal PNRR, è dedicato alla digitalizzazione dei depositi dei musei italiani in cui sono custoditi un numero elevatissimo di beni culturali non esposti nelle sale. Si tratta di beni talvolta visibili in mostre temporanee o specifici eventi che non ne consentono una continua fruibilità al pubblico. A beneficiare di tale investimento sono musei e aree archeologiche di tutto il territorio italiano, con particolare attenzione al Sud Italia destinatario del 40% dei fondi, per la messa a disposizione di oltre 600 mila oggetti digitalizzati. Altro materiale di interesse nel progetto sono le opere grafiche (disegni, stampe, bozzetti e matrici incise), conservate negli istituti afferenti alla Direzione Generale Musei quale il Gabinetto Disegni e Stampe delle Gallerie degli Uffizi.

Sotto il profilo normativo, il PND prevede cinque linee guida che offrono un quadro operativo multilivello d'ausilio nella pianificazione e messa in opera di tali progetti⁸⁶ e, come tale, si distinguono in:

- **Linee guida per i processi di digitalizzazione**⁸⁷. Costituiscono la base teorica e tecnica per la definizione delle metodologie per la creazione, metadattazione e archiviazione degli oggetti digitali. Forniscono anche indicazioni procedurali e amministrative indispensabili, nonché economiche, per la realizzazione dei progetti e in particolare nelle fasi iniziali e di collaudo.

⁸⁴ Istituto Centrale per la Digitalizzazione del Patrimonio Culturale – Digital Library, “AVVISI | NOTIZIE | PNRR PNRR, digitalizzazione del patrimonio culturale: online la procedura di gara da 16,9 milioni di euro per la categoria ‘Archivi fotografici: positivi, negativi, unicum e disegni’,” ultimo accesso 27 Gennaio 2024, <https://digitalibrary.cultura.gov.it/notizie/pnrr-digitalizzazione-del-patrimonio-culturale-online-la-procedura-di-gara-da-169-milioni-di-euro-per-la-categoria-archivi-fotografici-positivi-negativi-unicum-e-disegni/>.

⁸⁵ Istituto Centrale per la Digitalizzazione del Patrimonio Culturale – Digital Library, “PNRR, digitalizzazione del patrimonio culturale: online la procedura di gara da 18,6 milioni di euro per la categoria ‘Oggetti museali: beni di deposito (storico-artistici, archeologici) e grafici’,” ultimo accesso 27 Gennaio 2024, <https://digitalibrary.cultura.gov.it/notizie/pnrr-digitalizzazione-del-patrimonio-culturale-online-la-procedura-di-gara-da-186-milioni-di-euro-per-la-categoria-oggetti-museali-beni-di-deposito-storico-artistici-archeologici-e-gra/>.

⁸⁶ Istituto Centrale per la Digitalizzazione del Patrimonio Culturale, “Linee guida,” <https://digitalibrary.cultura.gov.it/linee-guida/>. Ultimo accesso 7 Gennaio 2024.

⁸⁷ Istituto Centrale per la Digitalizzazione del Patrimonio Culturale, *Linee guida per i processi di digitalizzazione 2022-2023* (2022), <https://docs.italia.it/italia/icdp/icdp-pnd-digitalizzazione-docs/it/v1.0-giugno-2022/index.html>.

- **Linee guida per la redazione del piano di gestione dei dati**⁸⁸. Scopo di tale documento è offrire le informazioni necessarie alla redazione del *Data Management Plan* (DMP) per i progetti di digitalizzazione e della gestione di banche dati sui beni culturali. La redazione del DMP è necessaria ai fini di garantire un’adeguata programmazione dei progetti documentandone le modalità di creazione, raccolta, gestione, conservazione e valorizzazione delle informazioni.
- **Linee guida per l’acquisizione, la circolazione e il riuso delle riproduzioni dei beni culturali in ambiente digitale**⁸⁹. Nell’ambito dei beni culturali è necessario garantire trasparenza nelle modalità e obiettivi per i quali una risorsa viene acquisita e impiegata, a uso commerciale e non. L’obiettivo di tali linee guida è dunque quello di guidare utenti e detentori dei dati – musei, archivi, biblioteche etc. – nella gestione del patrimonio culturale attraverso la districata normativa in vigore con il D.Lgs 42/2004 “Codice dei beni culturali e del paesaggio”. Tuttavia, occorre specificare che la pubblicazione delle linee guida è antecedente all’emanazione del DM 161/2023 “Linee guida per la determinazione degli importi minimi dei canoni e dei corrispettivi per la concessione d’uso dei beni in consegna agli istituti e luoghi della cultura statali” che, come è stato sottolineato nel comunicato delle diverse Associazioni e categorie di professionisti del settore culturale «il fatto che il nuovo DM 161 annoveri il rilascio e l’uso delle riproduzioni di beni culturali tra gli usi soggetti a concessione non appare coerente con il dettato delle precedenti Linee Guida cui abbiamo contribuito»⁹⁰.
- **Linee guida per la classificazione di prodotti e servizi digitali, processi e modelli di gestione**⁹¹. Il documento in questione si pone l’obiettivo di individuare e descrivere le metodologie per la classificazione dei servizi e prodotti oggetto del PND. La finalità delle linee guida è quella di inquadrare la funzione di ciascun servizio e prodotto all’interno del processo, delineandone gli aspetti positivi e negativi e le possibili ricadute sul patrimonio culturale.

⁸⁸ Istituto Centrale per la Digitalizzazione del Patrimonio Culturale, *Linee guida per la redazione del piano di gestione dei dati* (2022), <https://docs.italia.it/italia/icdp/icdp-pnd-dmp-docs/it/v1.0-giugno-2022/index.html>.

⁸⁹ Istituto Centrale per la Digitalizzazione del Patrimonio Culturale, *Linee guida per l’acquisizione, la circolazione e il riuso delle riproduzioni dei beni culturali in ambiente digitale* (2022), <https://docs.italia.it/italia/icdp/icdp-pnd-circolazione-riuso-docs/it/v1.0-giugno-2022/index.html>.

⁹⁰ Associazione Italiana Biblioteche, “Osservazioni sul DM 11 aprile 2023, n. 161,” <https://www.aib.it/notizie/osservazioni-sul-dm-11-aprile-2023-n-161/>. Ultimo accesso 7 Gennaio 2023.

⁹¹ Istituto Centrale per la Digitalizzazione del Patrimonio Culturale, *Linee guida per la classificazione di prodotti e servizi digitali, processi e modelli di gestione* (2022), <https://docs.italia.it/italia/icdp/icdp-pnd-servizi-docs/it/v1.0-giugno-2022/index.html>.

- **Introduzione alla metodologia per la valutazione della maturità digitale degli istituti culturali**⁹². Le ultime linee guida sono dirette a descrivere gli strumenti e metodologie da adottare per una concreta valutazione in merito alla maturità digitale degli istituti culturali. Questa è da intendersi come l'insieme degli obiettivi perseguiti, dei risultati ottenuti e delle attività in essere degli istituti su cui fondare il progetto di digitalizzazione. Nelle linee guida sono descritti i principali modelli di *digital maturity assessment*, in particolare quelli adottabili nel contesto culturale, con l'obiettivo ultimo di implementare uno strumento rispondente alle esigenze a livello centrale e periferico.

Le linee guida per la digitalizzazione del patrimonio culturale, pur non entrando nel dettaglio in merito all'applicazione delle tecnologie – non è difatti questo l'obiettivo dichiarato – descrivono il contesto entro cui un progetto di digitalizzazione deve svilupparsi individuando, a tal fine, quattro macro obiettivi:

1. conservazione degli originali;
2. fruizione e valorizzazione dei beni;
3. studio del patrimonio;
4. recupero di campagne di digitalizzazione pregresse.

È quantomai necessario inserire la conservazione tra gli obiettivi di un processo che, in taluni casi, risulta essere una lama a doppio taglio dove, alla finalità di promuovere e mettere a disposizione su più larga scala il patrimonio culturale, si contrappone il pretesto per la distruzione del bene materiale. È il caso della documentazione analogica prodotta e conservata da enti pubblici e privati per i quali, in un'ottica di ampliamento dello spazio fisico, la digitalizzazione è percepita come un'opportunità di eliminazione degli originali analogici. Si tratta di un mal costume che non tiene conto della complessità della situazione dal punto di vista storico-culturale, legale e tecnico. Pur se il valore probatorio può essere garantito, nella copia informatica, mediante apposita certificazione di processo⁹³, ciò non deve costituire il presupposto per l'eliminazione dell'originale. In tale circostanza, è alto il rischio di costituire una società fondata su copie, pur probatorie, manipolabili e privati della possibilità di effettuare raffronti con gli originali che, per loro natura, costituiscono la fonte certa delle informazioni. È soprattutto in prospettiva di una conservazione a lungo termine che si manifesta tale

⁹² Istituto Centrale per la Digitalizzazione del Patrimonio Culturale, *Introduzione alla metodologia per la valutazione della maturità digitale degli istituti culturali* (2022), <https://docs.italia.it/italia/icdp/icdp-pnd-maturita-docs/it/v1.0-giugno-2022/index.html>.

⁹³ D.lgs. 82/2005 *Codice dell'Amministrazione Digitale*, art. 22, <https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:decreto.legislativo:2005-03-07;82>. Per un dettaglio dei requisiti tecnici richiesti nella certificazione di processo Cfr. Agenzia per l'Italia Digitale (AGID), *Linee Guida sulla Formazione, Gestione e Conservazione dei Documenti Digitali* (2021), Allegato 3 "Certificazione di Processo," https://www.agid.gov.it/sites/default/files/repository_files/allegato_3_certificazione_di_processo.pdf.

necessità per cui, disporre esclusivamente delle copie informatiche, non rappresenta una soluzione ideale. In merito, il CAD (Codice dell'Amministrazione Digitale) stabilisce all'art.22, comma 5 che «con decreto del Presidente del Consiglio dei ministri possono essere individuate particolari tipologie di documenti analogici originali unici per le quali, in ragione di esigenze di natura pubblicistica, permane l'obbligo della conservazione dell'originale analogico»⁹⁴. Il testo qui riportato evidenzia l'importanza della forma originale analogica e, come tale, la necessità di salvaguardarla e conservarla. Basti pensare, infatti, alla fragilità di taluni supporti digitali e la frequenza con la quale vengono aggiornati o realizzati nuovi formati rendendo, di fatto, indispensabile effettuare continue operazioni di riversamento per evitare il rischio di obsolescenza e perdita delle informazioni. Si prenda ad esempio il recupero di un documento su un supporto ormai obsoleto quale un floppy disk. In tale situazione, ammettendo che il floppy disk non sia stato danneggiato in alcun modo, è necessario ricorrere ad un apposito lettore e, successivamente, di un software per la decodifica e lettura del formato del documento. Tuttavia, può accadere che, la disponibilità del lettore e del software per la lettura non sempre offra la certezza del pieno recupero delle informazioni desiderate. Difatti, occorre tenere conto anche di altri fattori quali, ad esempio, l'integrità del formato impiegato per la codifica del documento che può presentare un livello basso di robustezza ed essere facilmente soggetto a corruzioni e, di conseguenza, risultare illeggibile.

Sotto questo aspetto, invece, l'analogico presenta una maggiore certezza, a fronte di adeguati processi conservativi, consentendo una lettura dei documenti nel corso degli anni. Ne sono un esempio i papiri egiziani giunti sino ad oggi in buone condizioni, o le pergamene custodite dalla Camera dei Lord del Parlamento Inglese ancora impiegati – sino a qualche anno fa⁹⁵ – per la redazione dei documenti di maggior rilevanza del Regno Unito.

Anche da un punto di vista storico-culturale la conservazione degli originali si rende necessaria per preservarne l'autentico valore. Il digitale, pur avendo l'immenso vantaggio di promuovere oltre ogni confine i beni culturali dando loro maggiore visibilità, non trasferisce a questi il valore degli originali. Sono infatti questi ultimi, a costituire in molti casi un tratto distintivo dei paesi nonché l'attrazione principale dei viaggi turistici e simbolo di luoghi come, ad esempio, la *Mona Lisa* per il Museo del Louvre a Parigi o la *Primavera* di Botticelli per la Galleria degli Uffizi a Firenze. Si tratta in entrambi i casi di opere d'arte note in tutto il mondo che, seppur siano oramai ben visionabili attraverso la rete, continuano a suscitare grande interesse ed essere tra le principali attrazione nei rispettivi musei.

⁹⁴ D.lgs. 82/2005 *Codice dell'Amministrazione Digitale*, art. 22, comma 5, <https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:decreto.legislativo:2005-03-07;82>.

⁹⁵ "Gb, la Camera dei Lord rinuncerà ai documenti scritti su pelle di vitello," *La Repubblica*, 11 Febbraio 2016, https://www.repubblica.it/esteri/2016/02/11/foto/gb_la_camera_dei_lord_rinuncera_ai_documenti_scritti_su_pelle_di_vitello-133189845/1/.

Quanto detto non è esclusivo dell'ambito museale ma investe l'intero settore culturale. In ambito archeologico ad esempio, l'applicazione delle diverse tecnologie ha permesso la ricostruzione virtuale, in alcuni casi anche diacronica⁹⁶, dei siti offrendo agli utenti la possibilità di esplorare i luoghi attraverso un pc o persino uno smartphone. In Italia, il Ministero della Cultura ha promosso la realizzazione del *Gran virtual tour*, un progetto finalizzato alla scoperta del patrimonio culturale italiano in tour virtuale mediante la collaborazione con partner tecnologici quali Google Arts & Culture e ItalyArt⁹⁷. In campo internazionale, ad esempio, è possibile visitare le antiche rovine del tempio di Baalbek o delle Terme di Caracalla imperiale mediante una ricostruzione 3D al III sec. D.C.⁹⁸ direttamente da un dispositivo mobile.

La digitalizzazione offre dunque infinite possibilità per usufruire del patrimonio culturale, oltre che per curiosità o vago interesse, anche per scopi scientifici fornendo a studiosi e ricercatori un ambiente immersivo in cui lavorare. Ciò consente, ad esempio, di effettuare simulazioni circa il deterioramento o la resistenza a sollecitazioni di beni fragili⁹⁹, studiare e tradurre iscrizioni su manufatti fragili¹⁰⁰ o per approfondire dettagli di vario genere. L'*International Image Interoperability Framework* (IIIF) costituisce, a tal proposito, uno dei progetti di maggior successo per la digitalizzazione e fruizione del patrimonio culturale a livello internazionale. L'obiettivo del IIIF è la realizzazione di un ambiente digitale, interamente gratuito, nel quale gli attori coinvolti – utenti, informatici, curatori di beni culturali – hanno la possibilità di collaborare tra loro perseguendo il medesimo fine, ovvero la promozione dei beni culturali. Grazie agli strumenti e alle caratteristiche di cui è dotato e per l'impatto che ha sulla fruizione, a livello internazionale, del patrimonio culturale di vario genere (documentale, museale, artistico etc.), il IIIF è tra le tecnologie consigliate all'interno del già citato Piano nazionale di digitalizzazione del patrimonio culturale¹⁰¹ per l'esposizione e la gestione dei dati in accordo ai principi FAIR (*Findable, Accessible, Interoperable, Reusable*)¹⁰².

⁹⁶ Christos Kyrlitsias et al., "Corrigendum: A Virtual Tour of a Hardly Accessible Archaeological Site: The Effect of Immersive Virtual Reality in User Experience, Learning and Attitude Change," *Frontiers in Computer Science* 3, (2021). DOI: 10.3389/fcomp.2021.697259.

⁹⁷ Ministero della Cultura, "Grand virtual tour," <https://www.beniculturali.it/virtualtour>. Ultimo accesso 7 Gennaio 2024.

⁹⁸ FlyerZone, "Yorescape," <https://yorescape.com/>. Ultimo accesso 7 Gennaio 2024.

⁹⁹ Yoonjo Choi, Yun-Jung Yang, and Hong-Gyoo Sohn, "Resilient cultural heritage through digital cultural heritage cube: Two cases in South Korea," *Journal of Cultural Heritage* 48, (2021). DOI: 10.1016/j.culher.2021.01.007.

¹⁰⁰ Javier Andreu, and Pablo Serrano, "Contributions of the digital photogrammetry and 3D modelling of Roman inscriptions to the reading of damaged tituli: An example from the Hispania Tarraconensis (Castiliscar, Saragossa)," *Digital Applications in Archaeology and Cultural Heritage* 12, (2021). DOI: 10.1016/j.daach.2019.e00091.

¹⁰¹ Istituto Centrale per la Digitalizzazione del Patrimonio Culturale, *Piano nazionale per la digitalizzazione del patrimonio culturale 2022-2023* (2022), https://digitallibrary.cultura.gov.it/wp-content/uploads/2023/04/PND_V1_1_2023_v2.pdf.

¹⁰² Mark D. Wilkinson, "The FAIR Guiding Principles for scientific data management and stewardship," *Scientific Data* 3, (2016). DOI: 10.1038/sdata.2016.18.

Le considerazioni elaborate in merito alle sfide legate alla digitalizzazione hanno lo scopo di evidenziare la peculiarità dei beni culturali su cui le stesse società si riconoscono e trovano fondamento come evidenziato nella “Convenzione per la salvaguardia del patrimonio culturale immateriale”, edita dall’UNESCO nel 2003, in cui si legge

«Questo patrimonio culturale immateriale, trasmesso di generazione in generazione, è costantemente ricreato dalle comunità e dai gruppi in risposta al loro ambiente, alla loro interazione con la natura e alla loro storia e dà loro un senso d’identità e di continuità, promuovendo in tal modo il rispetto per la diversità culturale e la creatività umana»¹⁰³.

Da ciò emerge, in conclusione, la necessità di compiere ogni sforzo possibile verso la salvaguardia e integrità del patrimonio culturale senza commettere l’errore di considerare, come pari o sostitutivi agli originali, le copie altrimenti il rischio è di fondare identità, conoscenze e memorie soltanto sul virtuale. A tal riguardo costituisce un esempio quanto accaduto a Singapore dove il Governo aveva annunciato, nel 2011, di procedere alla rimozione del cimitero di Bukit Brown¹⁰⁴, considerato dalla popolazione come un importante patrimonio storico-culturale, per motivi urbanistici. Dal completamento dei lavori, avvenuto nel 2018, ciò che è rimasto di tale patrimonio sono i documenti digitalizzati fruibili nel portale dedicato¹⁰⁵, i soli ormai a mantenere vivo la memoria e il valore socio-culturale del luogo.

¹⁰³ United Nations for Education, Science and Culture Organization (UNESCO), *Convenzione per la salvaguardia del patrimonio culturale immateriale* (2003), art. 2, <https://www.lazio.beniculturali.it/wp-content/uploads/2022/05/ConvenzionePatrimonioImmateriale2003-ITA.pdf>.

¹⁰⁴ Ocon, David. “Digitalising endangered cultural heritage in Southeast Asian cities: preserving or replacing?,” *International Journal of Heritage Studies* 27, n. 10 (2021). DOI: 10.1080/13527258.2021.1883711.

¹⁰⁵ “Bukit Brown project,” <https://www.bukitbrown.info/about.php>. Ultimo accesso 7 Gennaio 2024.

3. Le metodologie di descrizione standardizzata di immagini come strumento di diffusione e valorizzazione del patrimonio culturale

Il modello International Image Interoperability Framework

All'interno del settore culturale, le immagini costituiscono uno dei mezzi principali per la disseminazione del patrimonio. Non a caso, durante i primi mesi della pandemia da Covid-19, si è assistito ad un incremento esponenziale di progetti aventi per oggetto la messa a disposizione via web di dipinti, manoscritti, fondi archivistici e così via. Le immagini sono dunque diventate il *medium* attraverso cui rendere documenti scritti, opere d'arte, nonché oggetti fisici, disponibili e accessibili a chiunque via Web. La digitalizzazione, così intesa, è dunque uno strumento in grado di ampliare «*le modalità di fruizione del patrimonio culturale, promuovendo e garantendo in modo attivo lo sviluppo della cultura [...]*»¹⁰⁶.

Uno degli elementi da valutare con attenzione nei progetti di digitalizzazione, a prescindere dall'obiettivo finale, è la scelta del formato elettronico che si intende adoperare. Come indicato nel Capitolo 3 delle stesse Linee Guida per la Digitalizzazione del patrimonio culturale e nelle Linee Guida AGID sulla Formazione, gestione e conservazione dei documenti digitali al Capitolo 3.6¹⁰⁷, la scelta del formato va effettuata prendendo in considerazione i seguenti fattori:

- Interoperabilità tra i sistemi;
- conservazione a lungo termine
- apertura del formato
- standardizzazione (preferibilmente *de jure*) del formato.

A questi principi generali, si affiancano determinati elementi valutativi legati alla sostenibilità del singolo progetto come la risoluzione delle risorse e i costi di realizzazione e mantenimento. Questi due aspetti sono posti in stretta relazione fra loro e costituiscono un punto focale nei progetti di digitalizzazione. Spesso, infatti, le ambizioni a produrre risorse ad altissima risoluzione si scontrano con i requisiti per il mantenimento di una struttura che richiede elevato spazio di conservazione e memoria di calcolo. Al contrario, elaborare una soluzione meno sofisticata comporta il rischio di realizzare materiale non adeguato. In entrambi i casi, questi compromessi corrono il rischio di influire negativamente sulle modalità di disseminazione culturale e public engagement.

¹⁰⁶ Istituto Centrale per la Digitalizzazione del Patrimonio Culturale – Digital Library. *Linee guida per la digitalizzazione del patrimonio culturale v. 1.0.* (2022), p.6. <https://docs.italia.it/italia/icdp/icdp-pnd-digitalizzazione-docs/it/v1.0-giugno-2022/il-progetto-di-digitalizzazione/perche-scopi-e-obiettivi-della-digitalizzazione.html>.

¹⁰⁷ Agenzia per l'Italia Digitale (AGID). *Linee guida sulla formazione, gestione e conservazione dei documenti informatici* (2021). https://www.agid.gov.it/sites/default/files/repository_files/linee_guida_sul_documento_informatico.pdf.

Una parziale soluzione a questi problemi è rappresentata dall'International Image Interoperability Framework (IIIF), un insieme di standard aperti per la condivisione semplice e veloce di materiale ad alta risoluzione. Nato nel contesto bibliotecario come mezzo di sviluppo delle digital library, la tecnologia IIIF si basa sui principi FAIR con l'obiettivo di rendere le risorse digitali accessibili all'interno di un ambiente collaborativo e interoperabile. È infatti sulla collaborazione che il framework fonda il proprio successo, divenendo in brevissimo tempo una comunità di enti di ricerca, istituti culturali e utenti di ogni genere con l'obiettivo di mantenere costantemente aggiornata l'intera architettura. Non avendo questo elaborato pretese di esaurire attentamente l'architettura IIIF, si procederà ad illustrarne le caratteristiche principali che costituiscono oggetto di applicazione all'interno del presente progetto di ricerca.

Funzionamento del IIIF

Il modello IIIF si compone di sei differenti *Application Profile Interface(s) -API-* così organizzate¹⁰⁸:

1. **Image.** Detta le specifiche per un web service che restituisca una immagine attraverso il protocollo http(s).
2. **Presentation.** Un insieme di metadati per la definizione della struttura del modello IIIF. Il risultato è un file in formato json che consentirà la visualizzazione della risorsa informativa attraverso specifici viewer.
3. **Authentication.** Definisce il processo di accesso alla risorsa tramite un sistema di controllo degli accessi.
4. **Change Discovery.** Provvede le informazioni necessarie alla navigazione e all'uso delle risorse IIIF.
5. **Content Search.** Definisce le modalità e la struttura affinché su una risorsa IIIF possano essere effettuate operazioni di annotazioni testuali.
6. **Content State.** Definisce le modalità per l'inizializzazione di una risorsa IIIF o parte di essa.

L'architettura logica di queste API si basa sulle specifiche REST¹⁰⁹ che consentono ai server, attraverso richieste con protocollo HTTP, di fornire dati in formato JSON-LD, una specifica del linguaggio JSON conforme ai principi dei Linked Data.

In questo elaborato saranno prese dettagliatamente in esame le Image API e Presentation API che costituiscono il nucleo centrale e indispensabile per la realizzazione di un modello IIIF di base.

¹⁰⁸ International Image Interoperability Framework (IIIF), "API Specifications," ultimo accesso 23 Gennaio 2023, <https://iiif.io/api/>.

¹⁰⁹ Salarelli, Alberto, "International Image Interoperability Framework (IIIF) : una panoramica," *JLIS* 8, n. 1 (2017), p. 53. DOI: 10.4403/jlis.it-12090.

Image API

L'Image API è un web service responsabile per la trasmissione in rete di una immagine con protocollo HTTP(S). Questa attività viene elaborata da adeguati image server che, ricevendo in input la risorsa, sono in grado di processarla e trasmetterla in rete secondo le precise specifiche dell'API. Il risultato di questa operazione è la generazione di un URI che identifica la risorsa attraverso una sequenza ordinata di parametri, consentendo all'utente di accedere ad una specifica rappresentazione dell'immagine. I parametri di cui l'URI si compone sono: *schema*, *server*, *prefisso*, *identificativo*, *region*, *size*, *rotation*, *quality* e *format*. Ciascun parametro risponde ad una funzione precisa che regola un aspetto dell'immagine. La tabella ivi riportata descrive sinteticamente le funzioni di ciascuna proprietà fornendo per ognuna di esse un esempio di valorizzazione.

Nome	Descrizione	Esempio
schema	Indica il protocollo impiegato per la trasmissione dell'immagine in rete.	http:// https://
server	Il server che ospita il servizio di trasmissione e può includere il numero di porta di hosting.	localhost:8182/ serverhosting.com/
prefix	Indica il percorso sul server ospitante dove risiede il servizio IIF. Si tratta di un elemento opzionale ma che può essere d'aiuto nei casi in cui su un server vengano ospitati più servizi.	image-service/
identifier	corrisponde al nome con cui la risorsa è memorizzata.	Esempio_foto1.jpg/
region	È il primo degli elementi per il recupero dell'immagine. Nello specifico indica se restituire l'intera immagine o una parte di essa. In questo caso devono essere specificati la forma o i pixel specifici secondo l'ordine x,y,w,h.	full/ square/ 125,0,52,450/ pct: 42,7,40,60/ (ridimensionamento in percentuale)
size	Indica le dimensioni attraverso cui restituire l'immagine. Anche in questo caso la scelta può essere fatta tra l'intera immagine o una sua scala specificando quale o quali pixel e il grado di ridimensionamento.	max/ 225/ ,225/ 225,250/ ¹¹⁰
rotation	Indica il grado di rotazione con cui restituire l'immagine. Il valore deve essere espresso con un valore numerico	0/ 180/ !0/

¹¹⁰ Per la lista completa degli esempi per l'elemento *size* si rimanda alla guida ufficiale del IIF sull'Image API: <https://iif.io/api/image/3.0/#42-size> ultimo accesso 23 Gennaio 2023.

	da 0 a 360. L'immagine può essere restituita con effetto riflesso aggiungendo il simbolo "!" davanti al valore numerico.	!270/
quality	Indica in che qualità l'immagine debba essere restituita: a colori, in scala di grigio o bitonale. Il valore <i>default</i> restituisce la qualità standard dell'immagine caricata.	gray/ dafult/ color/ bitonal/
format	Esprime il formato con cui l'immagine viene trasmessa. La disposizione dei formati tra cui scegliere dipende anche dall'immagine server impiegato	jpg/ tif/ png/

Di seguito un esempio di URI secondo l'Image API:

http://127.0.0.1:8182/iiif/3/iris_getty.jpg/full/max/90/default.png

Nell'esempio riportato, lo schema impiegato è il protocollo HTTP, il server su cui l'immagine viene trasmessa è una porta specifica (8182) aperta tramite la macchina virtuale locale `/127.0.0.1:8182/`. Il prefisso utilizzato `/iiif/3/` indica che per la visualizzazione dell'immagine si sta impiegando la versione 3.0 dell'API Image. A questi elementi introduttivi, fanno seguito l'identificativo dell'immagine `iris_getty.jpg` e le modalità attraverso cui si sta richiedendo l'accesso all'immagine, ossia: intera immagine `/full/`, nelle massime dimensioni `/max/`, ruotata di 90° in senso orario `/90/`, nella sua qualità standard `/default/` e, infine, con il formato `png /png/`. L'ordine in cui i parametri sono organizzati all'interno dell'URI è dettato dalle specifiche dell'Image API e, come tale, non è possibile invertire i parametri della richiesta.

Di recente, questa API è stata aggiornata alla versione 3.0 senza pur continuando a mantenere attiva la precedente versione 2.x (comprende la versione 2.0, 2.1 e 2.1.1), deprecando invece la versione 1.0. Questo aggiornamento si è reso necessario per eliminare ambiguità semantiche nell'uso di alcuni valori come, ad esempio, *full* che nella versione 2.x può essere attribuito sia alla proprietà *size* che *region*. I cambiamenti prodotti nel passaggio all'attuale versione possono essere ricondotti, per la gran parte, a modifiche di ordine morfologico con la ridenominazione di alcune proprietà. Altre modifiche, invece, hanno riguardato il piano strutturale intervenendo direttamente su alcuni singoli parametri e sui valori ad essi associati come nel caso del valore *full* che diviene, per la proprietà *size*, *max*.¹¹¹

¹¹¹ La lista completa dei cambiamenti prodotti nell'aggiornamento dell'Image API è consultabile alla specifica pagina web del IIIF : <https://iiif.io/api/image/3.0/change-log/>, ultimo accesso 23 Gennaio 2023.

Presentation API

La struttura cardine del IIIF è rappresentata dalla Presentation API, basata su JSON-LD, con l'obiettivo di definire l'intero modello per la rappresentazione dell'oggetto digitale nell'ambiente IIIF. A differenza dell'Image API che offre la possibilità di trasformare e accedere all'oggetto digitale via web, la Presentation API ne definisce la struttura composta dei seguenti otto oggetti principali:

- Collection
- Manifest
- Canvas
- Range
- Annotation Page
- Annotation
- Content
- Annotation Collection

Questi oggetti sono posti gerarchicamente in relazione tra loro secondo quanto espresso dallo schema nella Figura 1. A ciascuno di essi sono associati un id e un elemento *type* in grado di identificare univocamente gli oggetti all'interno della struttura. Questi elementi fanno parte di una serie di metadati, o per usare la terminologia di riferimento "*properties*", definiti dal IIIF per la descrizione del modello.

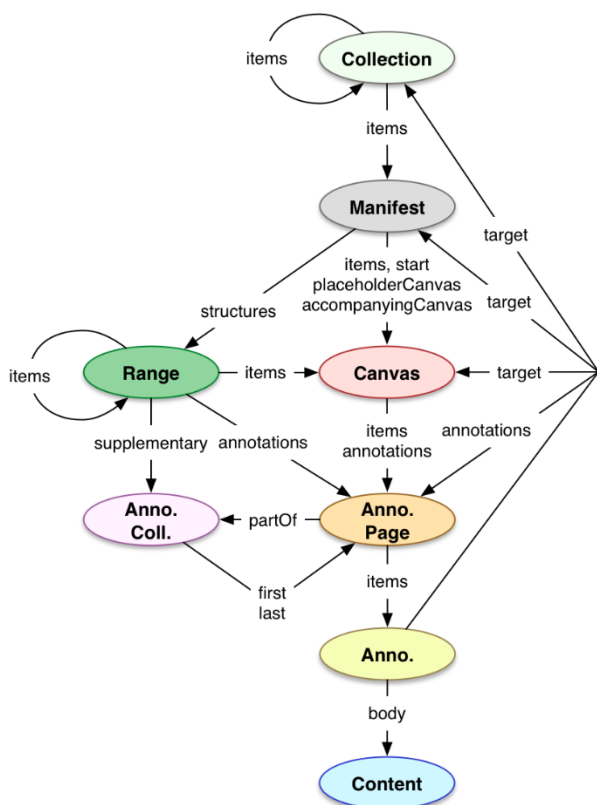


Figura 1: schema IIF Presentation API v. 3.0¹¹².

L'oggetto apicale del modello è il *Collection*, inteso come una lista ordinata di oggetti *Manifest* o di ulteriori *Collection*. Si tratta di un oggetto non obbligatorio, che può essere definito per organizzare più risorse digitali attribuendole un preciso ordine di visualizzazione. Il *Manifest* è l'elemento centrale all'interno di questa struttura e, come tale, è obbligatorio. Ha il compito di rappresentare la singola risorsa digitale nella sua globalità attraverso sotto diversi profili informativi in accordo ad uno specifico set di proprietà. Il *Canvas* si colloca all'interno del *Manifest* e svolge il ruolo di contenitore virtuale della risorsa digitale consentendo di attribuire alla risorsa una precisa rappresentazione. Al *Canvas* possono essere associati diverse risorse digitali come, ad esempio, immagini, testi e documenti audio-visivi ricorrendo agli *Annotation*, oggetti definiti secondo le specifiche del Web Annotation Data Model. Per la realizzazione dei collegamenti, all'interno dell'*Annotation* è necessario identificare univocamente la risorsa, ossia il *Content*, e il *Canvas* di destinazione. Questi oggetti devono essere identificati ricorrendo obbligatoriamente agli elementi *type* e *id* che, nel caso della risorsa contenuto, corrisponde all'URI ottenuto tramite Image API. Dal momento che ad ogni *Annotation* può corrispondere solamente una risorsa, è stato definito un oggetto lista, denominato *Annotation Page* con la funzione di ordinare al proprio interno i diversi *Annotation*. L'oggetto apicale nel caso delle "annotazioni" è costituito dall'*Annotation Collection*, al cui interno convergono i

¹¹² International Image Interoperability Framework (IIF), "Presentation API 3.0," ultimo accesso 23 Gennaio 2024, <https://iif.io/api/presentation/3.0/>.

diversi *Annotation Page*. Con la definizione di questo sovra-oggetto, l'obiettivo è quello di raggruppare *Annotation Page* per caratteristiche simili fra loro. Per chiarire il ruolo di questi oggetti, si prenda ad esempio il caso rappresentato nella Figura 2 di un manoscritto in lingua latina di cui sono state effettuate trascrizioni in diverse lingue.

Un *Annotation* può contenere la trascrizione di una singola pagina, o parte di essa mentre, l'insieme delle trascrizioni compiute in una determinata lingua, come nel caso dell'italiano (*Tras_it_tot*) sono raggruppate all'interno di un *Annotation Page*. Il ruolo svolto dall'*Annotation Collection* è quello di contenere al proprio interno le diverse trascrizioni suddivise per lingua (*Tras_it_tot*, *Tras_en_tot*...).

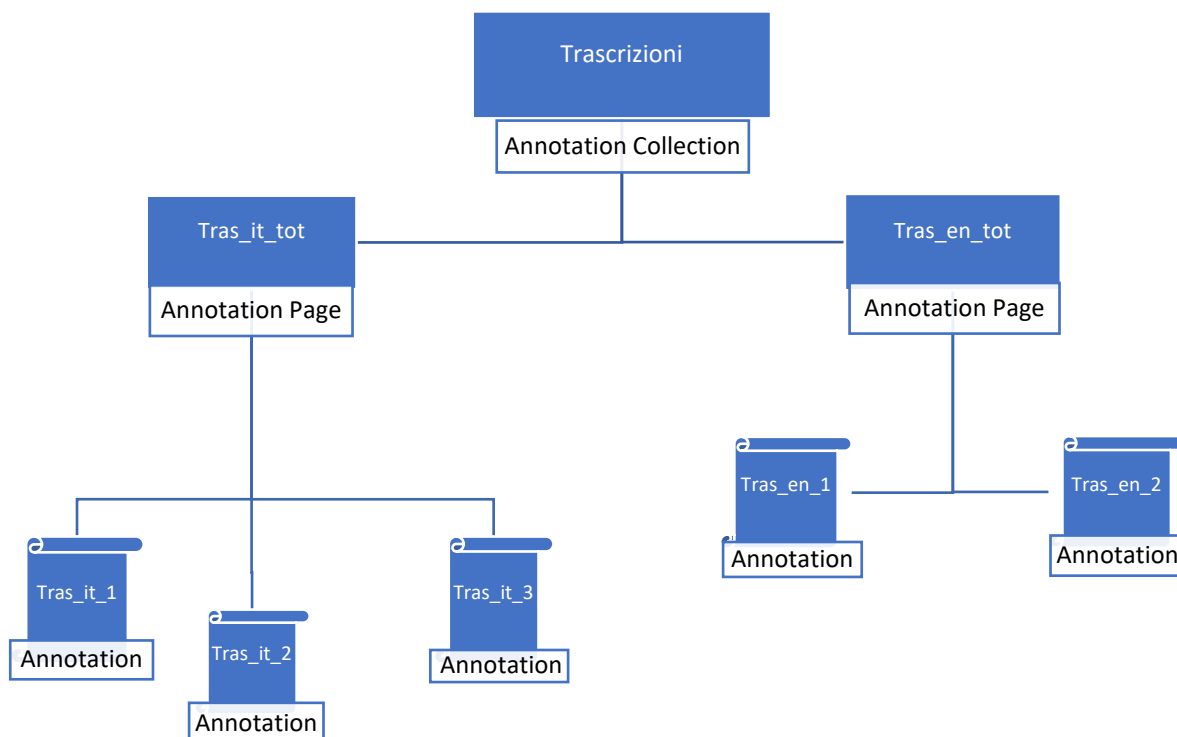


Figura 2: Esempio di applicazione degli oggetti *Annotation*.

L'ultimo oggetto che compone la struttura della Presentation API è *Range*, impiegato per definire uno specifico ordine di lettura e di visualizzazione dei *Canvas*, agevolando la navigazione dei contenuti all'interno del *Manifest* e del *Collection*.

Con l'intento di rispettare la specificità di ciascun oggetto che compone la struttura IIIF, sono state definite diverse proprietà distinte in: *descriptive*, *technical*, *structural* e *linking*. Ad eccezione di *id*, *type*, *item* e *label*, non è obbligatorio ricorrere ai metadati definiti nella struttura, seppure taluni siano espressamente raccomandati per una corretta descrizione e strutturazione delle risorse digitali. Le tabelle¹¹³ proposte di seguito offrono una sintesi dei metadati, suddivisi per categoria, indicando

¹¹³ International Image Interoperability Framework (IIIF), "Presentation API 3.0," ultimo accesso 23 Gennaio 2024, <https://iiif.io/api/presentation/3.0/>.

l'oggetto a quale possono riferirsi e il grado di obbligatorietà. Per facilitare la lettura delle tabelle, si tenga presente la legenda:

- M= obbligatorio
- R= raccomandato
- O= opzionale
- X= non permesso

	label	metadata	summary	id	navDate	language
<i>Collection</i>	M	R	R	M	O	X
<i>Manifest</i>	M	R	R	M	O	X
<i>Canvas</i>	R	O	O	M	O	X
<i>Annotation</i>	O	O	O	M	X	X
<i>AnnotationPage</i>	O	O	O	M	X	X
<i>Range</i>	R	O	O	M	O	X
<i>AnnotationCollection</i>	R	O	O	M	X	X
<i>Content Resources</i>	O	O	O	M	X	R

Tabella 2: Proprietà descrittive¹¹⁴.

	provider	thumbnail	accompanying Canvas	requiredStat ement	rights
<i>Collection</i>	R	R	O	O	O
<i>Manifest</i>	R	R	O	O	O
<i>Canvas</i>	O	O	O	O	O
<i>Annotation</i>	O	O	X	O	O
<i>AnnotationPage</i>	O	O	X	O	O
<i>Range</i>	O	O	O	O	O
<i>AnnotationCollection</i>	O	O	X	O	O
<i>Content Resources</i>	O	O	X	O	O

Tabella 3: Proprietà amministrative¹¹⁵.

	items	structures	Annotations
<i>Collection</i>	M	X	O
<i>Manifest</i>	M	O	O
<i>Canvas</i>	R	X	O

¹¹⁴ International Image Interoperability Framework (IIIF), "Presentation API 3.0," ultimo accesso 23 Gennaio 2024, <https://iiif.io/api/presentation/3.0/#31-descriptive-properties>.

¹¹⁵ International Image Interoperability Framework (IIIF), "Presentation API 3.0," ultimo accesso 23 Gennaio 2024, <https://iiif.io/api/presentation/3.0/>.

<i>Annotation</i>	X	X	X
<i>AnnotationPage</i>	R	X	X
<i>Range</i>	M	X	O
<i>AnnotationCollection</i>	X	X	X
<i>Content Resources</i>	X	X	O

Tabella 4: Proprietà di strutturali¹¹⁶.

	type	format	height	width	duration	behavior
<i>Collection</i>	M	X	X	X	X	O
<i>Manifest</i>	M	X	X	X	X	O
<i>Canvas</i>	M	X	O	O	O	O
<i>Annotation</i>	M	X	X	X	X	O
<i>AnnotationPage</i>	M	X	X	X	X	O
<i>Range</i>	M	X	X	X	X	O
<i>AnnotationCollection</i>	M	X	X	X	X	O
<i>Content Resources</i>	M	O	O	O	O	O

Tabella 5: Proprietà tecniche¹¹⁷.

Sotto il profilo sintattico, affinché un modello IIIF sia valido, è necessario che rispetti la codifica JSON-LD mentre, dal punto di vista strutturale, è richiesta la definizione di un *Manifest* senza obbligo di contenere al proprio interno un oggetto *Canvas*. Difatti, seppure la finalità sia la rappresentazione digitale della risorsa, è consentito il caso di un *Manifest* senza contenuto a cui sono associati solo dei metadati.

Con l'aggiornamento alla versione 3.0, la Presentation API ammette la possibilità di includere internamente anche oggetti audio-video, fino ad allora riferibili solo come collegamento esterno. L'aggiornamento di questa API porta con sé nuove ed interessanti potenzialità come la creazione di modelli sempre più complessi e interconnessi, offrendo maggiore contezza degli oggetti culturali tramite una rappresentazione più dettagliata e accurata. A questa nuova funzionalità hanno fatto seguito altri cambiamenti nell'aggiornamento dell'API. Non potendo esaurire dettagliatamente questo tema, in questa sede si procederà a descrivere schematicamente i principali cambiamenti occorsi dalla precedente versione 2.1.1 all'attuale 3.0 classificandoli in: cambiamenti sintattici e strutturali. In riferimento alla prima categoria, il principale cambiamento ha riguardato la sostituzione

¹¹⁶ International Image Interoperability Framework (IIIF), "Presentation API 3.0," ultimo accesso 23 Gennaio 2024, <https://iiif.io/api/presentation/3.0/#34-structural-properties>.

¹¹⁷ International Image Interoperability Framework (IIIF), "Presentation API 3.0," ultimo accesso 23 Gennaio 2024, <https://iiif.io/api/presentation/3.0/#32-technical-properties>.

dell'Open Annotation Model con il Web Annotation Data Model e, di conseguenza, la ridenominazione di proprietà e classi. Altre proprietà sono state rinominate talora per motivi tecnici, talaltra per garantire un uso semanticamente coerente della proprietà. Alcuni esempi di queste modifiche sono:

@id -> id

viewingHint -> behavior

attribution -> requiredStatement

Sotto il piano della strutturale, nel passaggio dalla versione 2.1 alla 3.0, il maggiore cambiamento occorso ha riguardato l'abolizione dell'oggetto Sequence a favore di Range con l'obiettivo di semplificare l'organizzazione delle risorse all'interno del modello.

Come molto spesso accade in ambiente informatico, gli aggiornamenti hanno avuto delle ripercussioni, in termini di compatibilità, sugli strumenti conformi al framework che necessitano di essere adeguati alle nuove specifiche per funzionare correttamente. In caso contrario, sarà impossibile impiegare, ad esempio, un image server per la realizzazione di una risorsa compatibile con le Image API 3.0.

Essendo il IIIF finalizzato alla condivisione delle risorse, un ruolo importante è svolto dai visualizzatori, strumenti software in grado di leggere il formato JSON-LD del modello prodotto tramite la Presentation API e darne visualizzazione secondo proprie specifiche. Si tratta di strumenti, salvo qualche eccezione ad opera di aziende che operano nel settore delle Digital Libraries, sviluppati da comunità di sviluppatori che basano il proprio operato sul principio dell'open access concedendo liberamente le specifiche degli strumenti. La stretta collaborazione tra sviluppatori e utenti, a rimarcare ulteriormente l'importanza rivestita dall'open access, costituisce un elemento di forte aggregazione all'interno di tutta la community IIIF, che avviene in forma di confronto diretto e immediato tramite piattaforme collaborative quali Slack¹¹⁸. Questo aspetto offre modo agli utenti di essere protagonisti attivi nei processi di sviluppo e aggiornamento degli strumenti.

La descrizione delle risorse secondo lo standard IIIF

Come si è detto, l'ambiente IIIF distingue al proprio interno quattro classi di metadati: descrittivi, strutturali, tecnici e di collegamento. Queste classi sono state definite con l'obiettivo di creare interoperabilità, all'interno del circuito IIIF il quale, per sua precipua natura, «*explicitly does not aim to provide metadata that would allow a search engine to index digital objects*»¹¹⁹ prediligendo un

¹¹⁸ Slack. <https://app.slack.com/>. All'interno del canale « IIIF » esistono diversi sotto gruppi specifici per singolo tema. Es. « 3d », « museums », « general » etc.

¹¹⁹ International Image Interoperability Framework (IIIF), "Presentation API 3.0," ultimo accesso 23 Gennaio 2023, <https://iiif.io/api/presentation/3.0/>.

approccio libero alla descrizione degli oggetti. Questa scelta, se da un lato lascia ampio margine decisionale agli utenti sui metadati da utilizzare, dall'altro si rivela un ostacolo per una descrizione uniforme e coerente delle risorse a cui i diversi standard di metadati cercano di ovviare nelle rispettive discipline. L'assenza di standardizzazione, in particolare sotto il profilo descrittivo, favorisce l'interazione tra utenti appartenenti a diversi campi culturali che, non trovandosi vincolati a schemi predefiniti, hanno modo di condividere la medesima struttura pur rispettando ciascuno le proprie procedure descrittive.

Con l'obiettivo di ottemperare alle diverse esigenze descrittive che intercorrono nei diversi domini GLAM (*galleries, libraries, archives and museum*), è stata definita la proprietà *metadata*, una lista composta di coppie *label-value* liberamente qualificabili. Per mantenere un alto grado di flessibilità, questa proprietà è stata strutturata per non imporre limiti derivanti da schemi predefiniti di metadati, salvo il rispetto della sola sintassi JSON-LD. Una caratteristica di questi elementi è il riferimento linguistico, vale a dire la possibilità di identificare la lingua impiegata nella valorizzazione del metadato tramite il codice identificativo dettato dalla ISO 639-1¹²⁰. Questa funzionalità consente agli utenti di specificare la lingua utilizzata per il proprio schema di metadati, ed eventualmente tradurlo. Tuttavia, uno schema così definito corre il rischio di essere solamente *human readable* e, di conseguenza, contrario alla definizione stessa di metadati come «*machine understandable information about web resources or other things*»¹²¹ e, più in generale come dati strutturati di dati. La compilazione libera e non controllata di metadati comporta il rischio di restituire informazioni che possono essere lette unicamente da umani, senza alcuna possibilità di interscambio tra diversi sistemi informativi. Per sopperire alla carenza di comprensione lato macchina, il modello IIF ammette il ricorso a risorse esterne attraverso la proprietà *seeAlso*. Come ogni altro elemento complesso all'interno della struttura, questa proprietà si configura come una lista di elementi atti a identificare in modo certo la risorsa esterna. A tal fine, è obbligatoria la valorizzazione degli elementi *id* e *type* ai quali è consigliato¹²² associare *label*, per dare alla risorsa un titolo interno al modello e *format* per specificarne la tipologia documentale e il formato. Il ricorso a questa proprietà si rende dunque necessario per associare una descrizione *machine readable* esterna alla risorsa IIF come, ad esempio, un record CIDOC per un oggetto museale, MARC nel contesto librario o, più genericamente, una descrizione con lo standard Dublin Core. Questa proprietà riveste un ruolo fondamentale per garantire

¹²⁰ ISO 639-1:2002. *Codes for the representation of names of languages – Part 1 : Alpha-2 code*.

¹²¹ W3C, "Metadata," ultimo accesso 23 Gennaio 2023, <https://www.w3.org/DesignIssues/Metadata.html>.

¹²² International Image Interoperability Framework (IIF), "Presentation API 3.0," ultimo accesso 23 Gennaio 2023, [https://iif.io/api/presentation/3.0/..](https://iif.io/api/presentation/3.0/)

all'interno del IIIF, una descrizione delle risorse conforme a standard specifici, permettendone «una migliore identificazione e, di conseguenza, un miglior reperimento»¹²³.

Insieme alle proprietà analizzate sinora, la struttura del IIIF integra un set di elementi descrittivi e amministrativi per l'identificazione delle principali caratteristiche della risorsa. Alcuni fra questi metadati sono:

- *summary*: elemento a testo libero per una breve descrizione della risorsa;
- *requiredStatement*: elemento usato per mappare i processi di attribuzione e provenienza della risorsa, è strutturato al proprio interno come una coppia *label-value* per specificare l'attributo del processo e il relativo valore;
- *rights*: identifica la licenza e i diritti sulla proprietà. Seppure l'elemento non imponga l'uso di un vocabolario controllato per la valorizzazione dell'elemento, si consiglia di inserire l'URI della licenza connessa alla risorsa come, ad esempio, nel caso delle creative commons:
`{ "rights": "http://creativecommons.org/licenses/by/4.0/" };`
- *navDate*: elemento da valorizzare secondo la codifica standard XSD *dateTime* literal per la rappresentazione di un valore temporale. Questo elemento ha l'obiettivo di collocare la risorsa entro un determinato arco temporale consentendone, in specifici casi, la visualizzazione tramite timeline;
- *provider*: elemento per la descrizione del soggetto che ha provveduto alla realizzazione o messa a disposizione del contenuto descritto. In particolare, questo elemento è strutturato come una lista JSON contenente al suo interno ulteriori sotto-elementi per l'identificazione univoca all'interno del record IIIF (*id*, *type*, *label*) e per il riferimento a risorse esterne (*homepage*, *logo*, *seeAlso*).

Questi elementi costituiscono il nucleo descrittivo per il quale il IIIF presenta una struttura articolata e complessa attuando un compromesso tra interoperabilità e libertà di scelta. Con un approccio diverso è stata impostata invece la sezione tecnica e strutturale che si compone di metadati definiti allo scopo di documentare univocamente la risorsa nel suo insieme. Tra questi, un ruolo centrale è svolto da *id* e *type* che consentono di identificare la risorsa nei diversi livelli di cui compone (*Collection*, *Manifest*, *Canvas* etc.). Questi elementi sono indispensabili per la corretta definizione della struttura del modello e propedeutici alla messa in visualizzazione dello stesso. Per definire adeguatamente ciascun oggetto e il livello in cui si colloca nel modello, è necessario valorizzare l'elemento *id* rispettando la gerarchica per evitare ambiguità strutturali.

¹²³ Mauro Guerrini, *Metadattazione*, (Milano: Editrice Bibliografica, 2022), p. 42. ISBN: 978-8893575195.

Ad esempio, supponendo di avere un modello strutturato così come nella Figura 3, con “manifest_opera” come oggetto apicale di *Manifest* contenente due oggetti *Canvas*, rispettivamente “canvas-1” e “canvas-2”, a cui sono associate delle immagini denominate “img-1” e “img-2” tramite la struttura di *Annotation Page* “anno-1”. In questa situazione si avranno i seguenti identificativi:

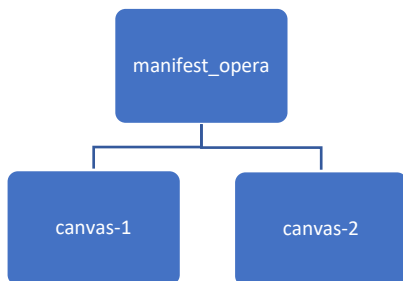


Figura 3: struttura di un modello IIIF

- manifest_opera: http://server_hosting.it/iiif/book/manifest_opera
- canvas-1: http://server_hosting.it/iiif/book/manifest_opera/canvas-1
- canvas-2: http://server_hosting.it/iiif/book/manifest_opera/canvas-2

A partire da questa struttura, è possibile procedere verso l’associazione delle risorse ai *Canvas* ricorrendo agli oggetti *Annotation* e *AnnotationPage*. Di conseguenza, riprendendo l’esempio precedente, si ipotizzi una struttura come quella proposta nella Figura 4:

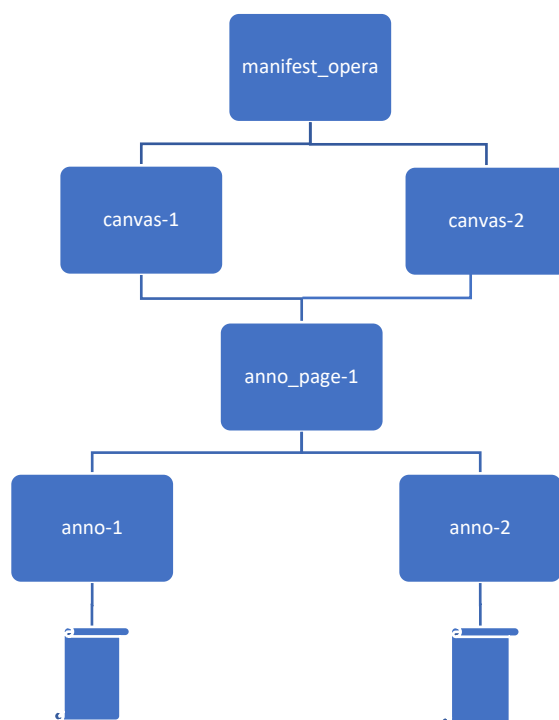


Figura 4: modello IIIF con immagini

In questo secondo esempio due risorse diverse, associate rispettivamente all’*Annotation* “anno-1” e “anno-2”, sono raggruppate all’interno di un unico *AnnotationPage* “anno_page-1”. Quest’ultimo,

quale oggetto indipendente, può essere collegato ad uno o più *Canvas* così come da grafico. Le possibili alternative in questo caso potranno essere:

- img-1: http://server_hosting.it/iiif/book/manifest_opera/canvas-1/anno_page-1/anno-1 (se l'immagine è associata al "canvas-1");
- img-2: http://server_hosting.it/iiif/book/manifest_opera/canvas-1/anno_page-1/anno-2 (se l'immagine è associata al "canvas-1");
- img-1: http://server_hosting.it/iiif/book/manifest_opera/canvas-2/anno_page-1/anno-1 (se l'immagine è associata al "canvas-2");
- img-2: http://server_hosting.it/iiif/book/manifest_opera/canvas-2/anno_page-1/anno-2 (se l'immagine è associata al "canvas-2").

Nei casi sopraelencati, sono ammessi diversi scenari in cui una immagine può essere associata ad uno o l'altro *Canvas* o, eventualmente, ad entrambi. Questo elemento viene precisato attraverso l'*id* del *Canvas* precisando in tal modo quale *Annotation* contenuto nella lista di *AnnotationPage* deve essere associato.

L'*id* è dunque l'elemento che identifica non solo la risorsa in sé, ma anche il livello in cui si colloca, contribuendo alla definizione della struttura gerarchica del modello IIIF.

L'esempio di seguito riportato fornisce una visione complessiva e generale di un modello IIIF di base composto da un *Manifest* e un *Canvas* a cui è associata una immagine statica.

```
{
  "@context": "http://iiif.io/api/presentation/3/context.json",
  "id": "https://gist.githubusercontent.com/martincritelli/manifest_1.json",
  "type": "Manifest",
  "label": {"it": ["Tesi Dottorato"]},
  "summary": {"it": ["Esempio di Manifest per Tesi Dottorato"]},
  "metadata": [{"label": {"en": ["creator"]}, "value": {"it": ["Martin Critelli"]}},
    {"label": {"it": ["data"]}, "value": {"it": ["2023"]}}],
  "items": [{"id": "https://gist.githubusercontent.com/martincritelli/manifest_1/canvas",
    "type": "Canvas",
    "label": {"it": ["Pagina-1"]},
    "width": 1024,
    "height": 720,
    "items": [{"id":
"https://gist.githubusercontent.com/martincritelli/manifest_1/canvas/annopage",
    "type": "AnnotationPage",
    "items": [{"id":
"https://gist.githubusercontent.com/martincritelli/manifest_1/canvas/annopage/annotation-1",
    "type": "Annotation",
    "motivation": "painting",
```

```

    "body": {
      "id": "https://localhost:8182/full/max/0/default.jpg",
      "type": "Image",
      "format": "image/jpeg",
      "width": 1024,
      "height": 720
    },
    "target": "https://gist.githubusercontent.com/martincritelli/manifest_1/canvas"
  ]
}
]
}
]
}
}

```

Casi applicativi

Sebbene abbia avuto origine nel campo bibliotecario per la rappresentazione ottimizzata via web di manoscritti e altro materiale librario, il IIIF ha conosciuto sin da subito una rapida espansione, con tempi e gradi diversi, anche negli altri domini GLAM, divenendo ben presto un modello di riferimento comune per l'intero settore culturale. Rispetto a biblioteche e musei, negli archivi il IIIF ha riscontrato maggiori difficoltà di espansione, nonostante sia un dominio quello archivistico caratterizzato da un'ampia eterogeneità documentale di elevato valore storico-culturale. Sono questi, infatti, elementi di grande interesse per la comunità IIIF e che pongono gli archivi storici in una posizione privilegiata per lo sviluppo di progetti di digitalizzazione ad ampio respiro che includano questa architettura e le connesse tecnologie ad accesso libero. Per inquadrare nel dettaglio la diffusione del framework nei diversi contesti culturali, in questa sezione si procederà a descrivere alcuni casi applicativi, delineando gli obiettivi e gli approcci adoperati in ciascuno di essi. Questo lavoro è stato condotto differenziando gli esempi per: biblioteche, musei e archivi.

IIIF e archivi

Il IIIF ha un notevole impatto nel settore archivistico pur non avendo raggiunto la medesima popolarità nei domini appena descritti. A livello internazionale, un punto di riferimento è costituito dalla British Library che ha in corso una graduale¹²⁴ realizzazione dell'architettura IIIF del proprio materiale archivistico e librario. Le diverse partnership instaurate con altre istituzioni e l'avvio di

¹²⁴ La lista dei fondi completamente disponibili è consultabile al sito <https://www.bl.uk/collection-guides/iiif>, ultimo accesso 23 Gennaio 2023,

specifici programmi¹²⁵, hanno permesso alla British Library di ampliare la rete di relazioni del proprio materiale agevolando agli utenti le attività di ricerca, accesso e analisi delle risorse documentarie. L'utente ha modo, dunque, di accedere alle risorse attraverso il catalogo generale della prestigiosa biblioteca dove le eventuali implementazioni in IIF sono segnalate dalla specifica icona dell'architettura tramite cui è possibile risalire al file descrittivo e al visualizzatore associato (Figura 5). Nella Figura 6 si riporta il caso di una corrispondenza tra il Questore provinciale di Mkumbira e il Questore distrettuale, il cui originale è attualmente conservato presso l'Archivio Nazionale del Malawi. Il documento è accessibile dall'Engangered Archives Programme (EAP) tramite il collegamento diretto al software Universal Viewer.



Figura 5: Schermata di accesso al documento¹²⁶.

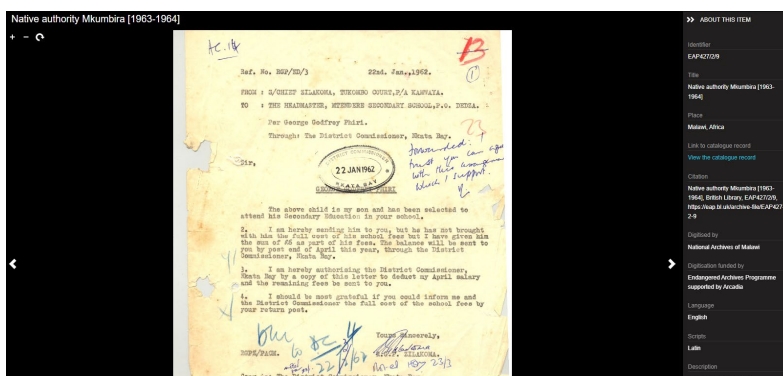


Figura 6: schermata di visualizzazione del documento tramite il software Universal Viewer accessibile direttamente dal sito dell'EAP.

¹²⁵ Il progetto di partnership della British Library ha coinvolto principalmente istituti del Medio Oriente e Sud Asia strettamente connessi con l'istituto britannico quali, ad esempio, la Qatar Digital Library e la National Library of Israel, oltre a specifici programmi di ricerca per la salvaguardia di patrimonio documentario a rischio come l'Endangered Archives Programme. Per maggiori informazioni circa i programmi di collaborazione della British Library si rimanda al sito ufficiale <https://www.bl.uk/international-engagement/partnerships>, ultimo accesso 23 Gennaio 2023.

¹²⁶ British Library, "Endangered Archives Programme," ultimo accesso 23 Gennaio 2023, <https://eap.bl.uk/archive-file/EAP427-2-9>.

Altra esperienza di successo a livello internazionale è quella di Biblissima, che si configura come un polo di aggregazione per diversi istituti quali la Bibliothèque nationale de France, la Biblioteca Nacional Digital de Portugal o ancora la Bayerische Staatsbibliothek. In questa sede, il progetto si distingue per l'ambizione di creare un prototipo in grado di organizzare e indicizzare i singoli *Manifest* attraverso i metadati in essi contenuti, normalizzare le stringhe di linguaggio e gli *authority file* ricorrendo a specifiche liste controllate come il VIAF, Wikidata, LoC etc., e sperimentare nuove tecniche di estrazione di metadati per la creazione automatica dei *Manifest* partendo dalle specifiche schede descrittive¹²⁷. All'interno di questo progetto il IIF si evolve, sul piano concettuale, sino ad assumere una nuova dimensione; non si configura solo come il risultato di un processo di digitalizzazione per il rapido accesso e scambio di oggetti digitali ma anche come preziosa fonte di informazioni da cui attingere per ampliare basi di conoscenza attraverso immagini, metadati e annotazioni testuali. All'interno del proprio database, la ricerca è stata organizzata per essere condotta sia all'interno del catalogo generale che filtrando per singole istituzioni. In aggiunta, Biblissima ha fatto ricorso all'API Change Discovery, le apposite specifiche per agevolare i processi di ricerca e accesso ai singoli oggetti digitali tra i diversi provider partecipanti. Partendo dai singoli oggetti messi a disposizione, l'utente è in grado di creare una propria collezione (Figura 7) da visualizzare nell'ambiente di Mirador 2 (Figura 8).

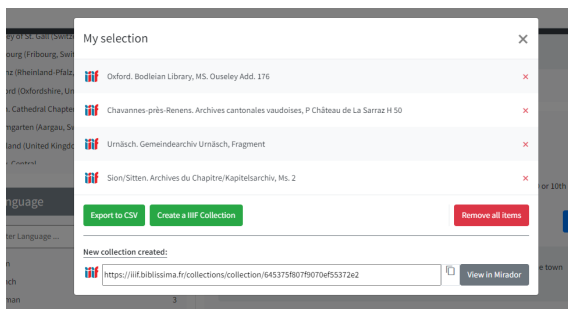


Figura 7: selezione e creazione della collezione IIF

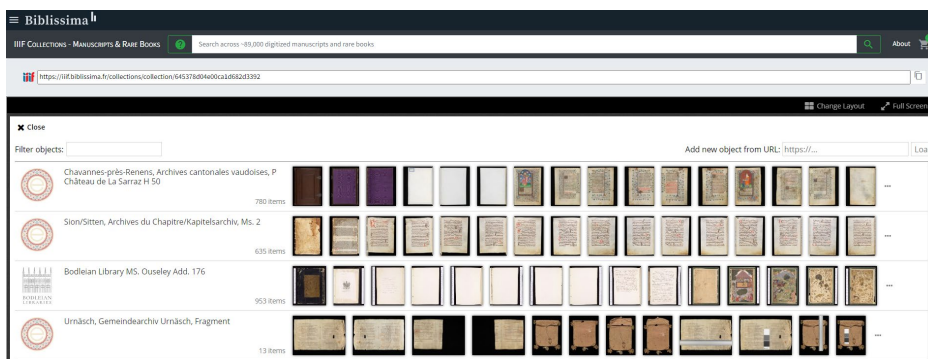


Figura 8: visualizzazione in Mirador 2 della collezione creata in precedenza.

¹²⁷ Biblissima, "Collections," ultimo accesso 23 Gennaio 2023, <https://iif.biblissima.fr/collections/about>.

Nel contesto nazionale sono stati realizzati alcuni significativi casi applicativi. Tra questi rientra il caso dell'Archivio storico dell'Università di Bologna¹²⁸ dove sono stati digitalizzati gran parte dei propri fondi includendo, tra i tanti, i verbali dei Consigli Accademici e di Senato, i fascicoli di Roberto Roveresi e Pier Paolo Pasolini. Così come in molti altri esempi analizzati, per questo progetto sono state impiegate le specifiche 2.0 delle API di Presentation e Image impostando una versione di base di Mirador 2 come visualizzatore predefinito.

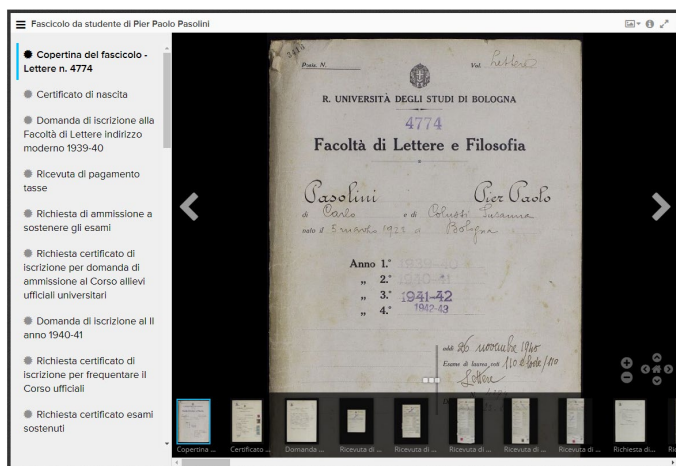


Figura 9: visualizzazione del fascicolo studente di Pier Paolo Pasolini¹²⁹.

Altri progetti che hanno suscitato grande interesse tra la comunità di studiosi e ricercatori, sono quelli promossi dall'Archivio storico della Gazzetta di Parma¹³⁰ e dalla Digital library dell'Archivio Centrale dello Stato¹³¹. Un elemento che accomuna questi progetti è, come nel caso delle Gallerie Estensi, l'implementazione tramite la piattaforma MLOL-DH la visualizzazione con la modalità *Stories*. Grazie anche alla reputazione e popolarità della Media Digital Library, questi progetti godono di una maggiore pubblicità e visibilità rispetto ad altri casi applicativi. Nel caso dell'Archivio storico della Gazzetta di Parma, la digitalizzazione è stata condotta con l'obiettivo di mettere in risalto il ruolo che lo storico quotidiano ha assunto all'interno della comunità locale. Per tale ragione, l'utente è guidato nella navigazione del Portale dell'Archivio della Gazzetta di Parma attraverso un percorso lineare suddiviso in tre fasi, o per meglio dire *storie*, corrispondenti ai momenti chiave del quotidiano:

¹²⁸ Archivio Storico Università di Bologna, "Patrimonio documentario," ultimo accesso 23 Gennaio 2023, <https://archiviostorico.unibo.it/it/patrimonio-documentario>.

¹²⁹ Archivio Storico Università di Bologna, "Fascicolo di Pier Paolo Pasolini in IIF," ultimo accesso 23 Gennaio 2023, https://archiviostorico.unibo.it/it/patrimonio-documentario/fascicolo-e-tesi-di-pier-paolo-pasolini/fascolo-da-studente-di-pier-paolopasolini?view=ImageView&manifest=https%3A%2F%2Farchiviostorico.unibo.it%2Fiiif%2F%2Fmanifest%2Farchiviostorico%2Ffascicolo_e_tesi_di_pier_paolo_pasolini%2Ffascicolo_pasolini.json&canvas=https%3A%2F%2Farchiviostorico.unibo.it%2Fiiif%2F%2Fcanvas%2Farchiviostorico%40fascicolo_e_tesi_di_pier_paolo_pasolini%40fascicolo_pasolini%402b8ccc5f6a848ca713a2bcffd5900ae032620d40-01.jp2.json.

¹³⁰ Media Library Online, "Archivio Storico Gazzetta di Parma," ultimo accesso 23 Gennaio 2023, <https://archiviogazzettadiparma.medialibrary.it/home/index.aspx>.

¹³¹ Archivio Centrale dello Stato, "Digital Library dell'ACS," ultimo accesso 23 Gennaio 2023, <https://acs.cultura.gov.it/e-online-la-digital-library-dellacs/>.

1. Zavattini. I primi anni della Gazzetta:1923-1927;
2. Quarto centenario della morte di Correggio. Le celebrazioni del 1934-35;
3. Alfredo Zerbini e l'Oltretorrente.

L'utente ha dunque la possibilità di navigare il contenuto digitalizzato attraverso la specifica piattaforma MLOL e visualizzarlo in ambiente IIF con la specifica versione di Mirador implementata.

Per concludere questa breve disamina su alcuni casi applicativi del IIF, si riporta l'esempio della Teca digitale dell'Archivio Centrale dello Stato (ACS)¹³², certamente tra i più ambiziosi progetti di digitalizzazione del nostro Paese in ragione della vasta mole di fondi e documenti conservata all'interno dell'ACS. Il progetto, che ha avuto avvio a luglio 2022, attualmente¹³³ conta decine di migliaia di documenti e fondi vari¹³⁴ digitalizzati e resi disponibili agli utenti tramite il già citato modulo *Stories* della Media Digital Library. Per le tipologie documentali digitalizzate e il valore storico-culturale che essi rivestono, il progetto Teca digitale dell'ACS costituisce un punto di riferimento in Italia per l'applicazione dell'architettura IIF al settore archivistico e, in generale, dei beni culturali. La messa a disposizione, nel caso illustrato, di materiale documentale storico degli organi centrali dello Stato in un ambiente facilmente gestibile e interoperabile, aiuta a comprendere le potenziali ricadute che l'applicazione del IIF agli archivi può avere sulla società. Si tratta dunque di uno strumento che, sapientemente utilizzato, offre agli archivisti l'opportunità di valorizzare e disseminare ad ampio raggio il materiale che hanno il compito di custodire.

Pur essendo evidente l'impatto che questi progetti hanno avuto nella valorizzazione della documentazione d'archivio, lo stesso non può dirsi sotto il piano tecnico trattandosi di applicazioni realizzate secondo le precedenti specifiche della versione 2 della Presentation API. Una scelta che riflette l'attuale stato di impiego del IIF nel contesto italiano, dove risultano essere sporadici i casi di applicazioni con le api aggiornate, ad eccezione di qualche rara eccezione come, ad esempio, il già citato progetto della Galleria Borghese.

IIF e biblioteche

Si tratta sicuramente del dominio in cui sono presenti i maggiori lavori di modellazione IIF. A livello internazionale i casi da prendere a esempio sono le collezioni create dalle Bodleian Library e Stanford Library, tra gli enti fondatori del IIF, mentre nel contesto nazionale figurano tra le diverse le Gallerie Estense e la Digital Vatican Library.

¹³² Archivio Centrale dello Stato, "Teca Digitale ACS," ultimo accesso 23 Maggio 2023, <https://tecadigitaleacs.cultura.gov.it/media/esplora>.

¹³³ Dati aggiornati a Maggio 2023.

¹³⁴ Ministero della Cultural, "Dal caos al cosmo," ultimo accesso 23 Gennaio 2024,

<https://www.beniculturali.it/evento/dal-caos-al-cosmo-la-nuova-teca-digitale-dellarchivio-centrale-dello-stato>.

La Bodleian Library dell'Università di Oxford ha disposto nell'ambiente IIIF l'intero materiale di cui è in possesso comprendente: manoscritti, incunaboli e materiale archivistico vario. Il materiale è reso disponibile di default in IIIF attraverso la Presentation API 2.1 con la possibilità di accedervi tramite la versione 3.0. L'interfaccia creato dalla Bodleian Library permette di navigare all'interno della collezione, oltre che per tipologia documentale, per mezzo dello status distinguendo tra “*fully digitalized*” e “*partially digitalized*”.

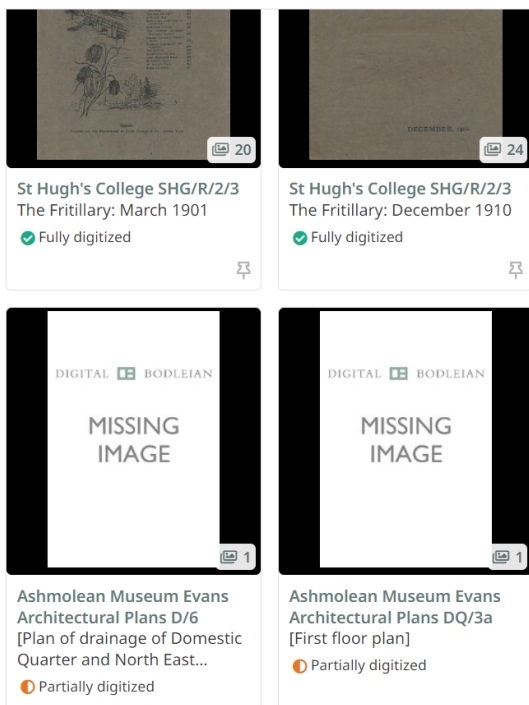


Figura 10: anteprima del catalogo della Bodleian Library¹³⁵.

Anche nel caso della Stanford Library si parla di una istituzione protagonista nella creazione dell'ambiente IIIF. Si tratta di un caso particolarmente interessante grazie al viewer open source Mirador sviluppato all'interno della stessa Stanford University, che consente la visualizzazione poliedrica delle risorse grazie all'integrazione con diversi plugin. Attualmente la nota biblioteca mette a disposizione la propria collezione secondo le specifiche 2.1 delle Presentation API con un collegamento diretto ad una versione personalizzata del viewer per consentire una navigazione semplice e immersiva del patrimonio librario in possesso.

Nel contesto italiano un ruolo di rilievo è assunto dalle Gallerie Estensi con l'Estense Digital Library (EDL) e della Biblioteca Apostolica Vaticana (BAV) con la Digital Vatican Library (DVL) accumulate, sotto il profilo concettuale, per la scelta di creare percorsi tematici con le immagini del proprio patrimonio librario e documentale. Di particolare interesse è la funzionalità denominata *IIIF Stories* che permette, come nel caso della EDL, di navigare la collezione attraverso un preciso filo

¹³⁵ Bodleian Library, “Modern manuscripts and archives,” ultimo accesso 23 Gennaio 2024, <https://digital.bodleian.ox.ac.uk/collections/modern-manuscripts-and-archives/?rows=100&page=7>.

conduttore. Questo servizio è realizzato dalla sezione *Digital Humanities* della Media Library Online (MLOL - DH)¹³⁶, tra i principali network di prestito digitale, a cui aderiscono diversi enti per la digitalizzazione del proprio patrimonio documentale e librario. La sezione DH fa ricorso all'infrastruttura IIF per la messa a disposizione del materiale digitalizzato, creando collegamenti tematici che mettano in risalto il contesto e rendano la navigazione fruibile come un percorso virtuale. La BAV sta portando avanti un progetto triennale in collaborazione con la Stanford University Libraries con l'obiettivo di digitalizzare in ambiente IIF il proprio patrimonio per renderlo fruibile e analizzabile attraverso un sistema di annotazioni, trascrizioni e comparazione di testi e immagini.¹³⁷ Per la realizzazione del modello sono state impiegate le API Image e Presentation alla versione 2.0, fruendo di una versione personalizzata di Mirador viewer in grado di sostenere il processo di annotazione degli utenti e memorizzarle attraverso il *Simple Annotation Server*¹³⁸.

IIF e musei

In ambito museale i casi applicativi sono innumerevoli e si differenziano fra loro per scopi e modalità di implementazione. In questa sede verranno illustrati brevemente i casi del Getty Museum, Natural History Museum di Londra e della Galleria Borghese di Roma che hanno portato avanti differenti modalità di realizzazione di architetture IIF. Il Getty Museum, già membro del Consorzio IIF e tra gli istituti più autorevoli nel settore museale internazionale, è stato tra i primi musei a presentare la propria collezione in IIF. Il progetto è stato realizzato con l'obiettivo di rendere il materiale liberamente accessibile e fruibile agli utenti ricorrendo alla digitalizzazione in alta risoluzione delle opere da diverse prospettive. L'esempio nella Figura 11¹³⁹ mostra la rappresentazione dell'*Iris* di Van Gogh, di cui sono state realizzate tre differenti prospettive: opera, opera incorniciata e retro del quadro. La realizzazione di un modello ad alta definizione, strutturato per contenere l'opera in tutti i suoi elementi costitutivi, offre un nuovo livello di approfondimento, consentendo a studiosi e ricercatori di accedere e analizzare la risorsa in remoto.

¹³⁶ Media Library Online, ultimo accesso 23 Gennaio 2024, <https://www.medialibrary.it/pagine/pagina.aspx?id=736>.

¹³⁷ Digital Vatican Library, "Overview," ultimo accesso 23 Gennaio 2024, <https://spotlight.vatlib.it/overview>.

¹³⁸ Robinson, Glen, "Simple Annotation Server," ultimo accesso 23 Gennaio 2024, <https://github.com/glenrobson/SimpleAnnotationServer>. Un server creato con l'obiettivo di conservare le annotazioni effettuate sulle risorse IIF. Attualmente supporta la versione 2.x delle Presentation API.

¹³⁹ Getty Museum, "Iris," ultimo accesso 23 Gennaio 2024, <https://www.getty.edu/art/collection/mirador/?manifest=https://media.getty.edu/iif/manifest/53be857e-41e8-4198-b45d-2e0f52d3051b>.

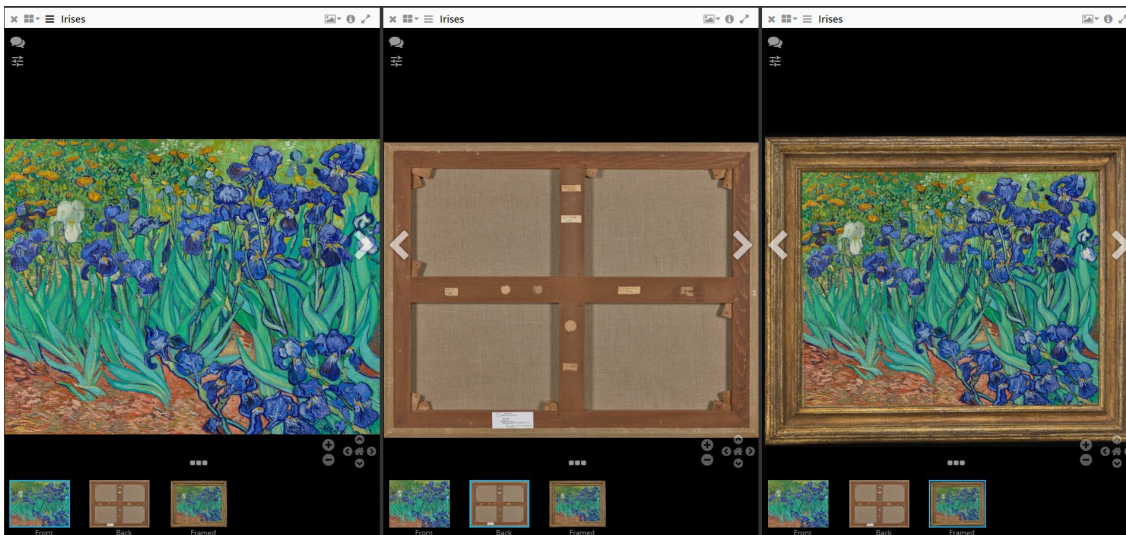


Figura 11: Visualizzazione dell'Iris di Van Gogh tramite il Mirador del Getty Museum

Attualmente la digitalizzazione è stata condotta impiegando la versione 2.1 delle API Image e Presentation mentre per la fase di visualizzazione si fa ricorso ad una versione di Mirador 2 che consente di effettuare annotazioni e altre funzionalità cromatiche sulle immagini.

Il Natural History Museum di Londra rappresenta uno dei primi casi di applicazione delle nuove specifiche a collezioni museali pur non avendo, allo stato attuale¹⁴⁰, reso disponibile in ambiente IIIF l'intero patrimonio. Questo istituto consente la visualizzazione delle risorse direttamente sul proprio sito tramite il plugin di *Openseadragon*¹⁴¹ che consente di effettuare profondi livelli di ingrandimento sulle immagini mantenendo un'alta risoluzione delle stesse. La scelta di questo software non limita tuttavia le potenzialità offerte dal IIIF, avendo l'utente la possibilità di impiegare altri visualizzatori ricorrendo all'identificativo del *Manifest* o del *Collection*.

L'ultimo caso preso in esame per la sezione museale del IIIF è quello della Galleria Borghese di Roma. Si tratta di un progetto di ricerca che ha coinvolto, oltre alla stessa galleria, il Venice Center for Digital and Public Humanities (VeDPH) dell'Università Ca' Foscari di Venezia e l'Istituto di Linguistica Computazionale del Consiglio Nazionale delle Ricerche (CNR-ILC) a cui ha preso parte il sottoscritto in qualità di Assegnista di ricerca. In questo progetto si è proceduto a «integrare le molteplici relazioni che intercorrono all'interno dell'intero catalogo di opere della Galleria Borghese nei differenti Manifest del IIIF». ¹⁴²

L'obiettivo è stato definito per valorizzare la componente collezionistica e spaziale del museo che, oltre al patrimonio monumentale, detiene un duplice carattere distintivo, caratterizzato da un ricco

¹⁴⁰ National History Museum of London, "Natural History Museum," ultimo accesso 23 Gennaio 2024, <https://data.nhm.ac.uk/>.

¹⁴¹ OpenCollective, "OpenSeadragon," ultimo accesso 23 Gennaio 2024, <https://openseadragon.github.io/>.

¹⁴² De Vincentis, Stefania, and Martin Critelli, "Mappare il museo in IIIF : Una combinazione di deep zoom e VR360 per la Galleria Borghese di Roma," *La memoria digitale: forme del testo e organizzazione della conoscenza. Atti del XII Convegno Annuale AIUCD*, (2023). ISBN – 978-88-942535-7-3.

apparato decorativo che fa da cornice alla collezione di sculture, dipinti e arredi. L'espedito adottato nella sperimentazione consiste nell'avvalersi di fotografie equi rettangolari per la rappresentazione lineare di ogni sala, che si configura a sua volta come una mappa tramite cui accedere alle singole opere. Il progetto è stato portato avanti impiegando le attuali specifiche 3.0 delle Image e Presentation API incorporando una versione di Mirador che permetta, tramite il sistema di annotazioni, di accedere alle opere a partire dall'immagine equi rettangolare della sala. La Figura 12 è un esempio del progetto strutturato attraverso l'elemento di *Collection* in grado di contenere al proprio interno l'intera struttura della sala e delle opere ad essa connessa.

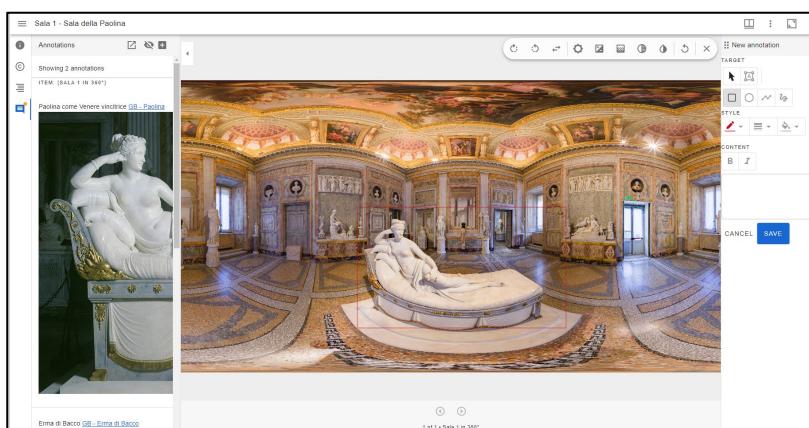


Figura 12: Visualizzazione della Sala 1 della Galleria Borghese con il viewer Mirador 3

Applicazione del IIIF al progetto

Come emerso dai casi d'uso citati, il IIIF si configura come uno dei possibili mezzi per la valorizzazione e disseminazione del patrimonio culturale. Per tale ragione, in questo progetto di ricerca, si è proceduto a definire un modello IIIF del campione di studio partendo dagli elementi ottenuti con l'applicazione degli strumenti di estrazione automatica. L'obiettivo di questa ulteriore fase applicativa è la valorizzazione del patrimonio culturale tramite raffigurazione degli oggetti culturali in un ambiente dinamico, interoperabile, e modellabile sotto il profilo tecnico. In particolare, si aspira a definire un modello utile alla valorizzazione del patrimonio culturale così come intesa all'interno del Codice dei Beni Culturali:

«attività dirette a promuovere la conoscenza del patrimonio culturale e ad assicurare le migliori condizioni di utilizzazione e fruizione pubblica del patrimonio stesso, anche da parte delle persone diversamente abili, al fine di promuovere lo sviluppo della cultura.»¹⁴³

A tale scopo, è stato adottato un approccio che consentisse la creazione di un modello di *Collection* uniforme, sotto il profilo semantico, facendo ricorso a standard di metadati. Per la realizzazione di

¹⁴³ D.Lgs. n. 42/2004, Codice dei beni culturali e del paesaggio, art.6.

questo obiettivo si è tenuto conto del variegato panorama di standard di metadati descrittivi esistenti distinguendoli tra quelli ad uso generico e specifici di dominio allo scopo di garantire una descrizione quanto più dettagliata possibile del patrimonio culturale. A tal fine, è stato necessario considerare i diversi punti di contatto tra le diverse tipologie documentali afferenti al dominio senza però trascurarne le dovute specificità. Trattandosi di risorse culturali visive, per mantenere un livello intermedio tra la generalità e la specificità nella descrizione delle risorse, sono stati adoperati in combinazione gli standard Dublin Core e Visual Resources Association Core (VRA Core). Questa scelta è stata adoperata per sfruttare, da un lato, l'eterogeneità e granularità dei metadati del Dublin Core e, dall'altro, la specificità di dominio del VRA Core con l'obiettivo di creare un *application profile*, ossia:

«a declaration of which metadata terms an organization, information resource, application, or user community uses in its metadata»¹⁴⁴

A tale proposito, un aspetto essenziale su cui porre l'attenzione concerne i *metadata crosswalk*, mappature tra elementi di standard differenti per garantire interoperabilità, sotto il profilo semantico e sintattico, tra i diversi schemi di metadati. Tipicamente, queste mappature sono effettuate ricorrendo a grafici o tabelle in grado di veicolare semanticamente un elemento dallo standard di riferimento verso uno diverso.¹⁴⁵

Allo stato attuale sono state definite sotto il profilo concettuale diverse mappature tra le quali, di particolare interesse per il numero di standard coinvolti, è quella del Getty Institute¹⁴⁶. Questo *crosswalk* è stato elaborato paragonando diversi standard di metadati quali Dublin Core, MARC, VRA Core, MODS, CDWA e il modello concettuale CIDOC CRM. Mappature tra metadati come quella definita dal Getty Institute, assumono un ruolo di rilievo nei diversi settori perché consentono alle singole istituzioni di riorganizzare le informazioni contenute nei propri sistemi per garantirne una maggiore interoperabilità. Queste “linee guida” costituiscono un valido strumento per catalogatori, esperti di dominio, ricercatori e studiosi, ognuno dei quali per le proprie finalità può trarre vantaggio dalla comparazione dei diversi standard. Ad esempio, si pensi al confronto tra Dublin Core e la famiglia MARC nel dominio librario; nonostante quest'ultimo si configuri come uno standard di dominio, i cui metadati sono stati definiti appositamente per la descrizione esaustiva del materiale bibliotecario, negli ultimi anni si sta prediligendo¹⁴⁷ l'uso del Dublin Core il cui semplice schema,

¹⁴⁴ Tom BAKER, “DCMI Usage Board review of application profiles,” ultimo accesso 23 Gennaio 2024, <https://www.dublincore.org/specifications/dublin-core/application-profile-review/>.

¹⁴⁵ Chan, Lois Mai, and Marcia Zheng, “Metadata Interoperability and Standardization – A Study of Methodology, Part I: Achieving Interoperability at the Schema Level,” *D-Lib Magazine* 12, n. 6 (2006), p. 8. DOI:10.1045/june2006-zeng.

¹⁴⁶ Getty Institute, “Metadata Standards Crosswalk,” ultimo accesso 23 Gennaio 2024, https://www.getty.edu/research/publications/electronic_publications/intrometadata/crosswalks.html.

¹⁴⁷ Eccezione fatta per i sistemi bibliotecari che rimangono ancora costruiti sullo standard UNIMARC.

con l'ausilio di liste di termini controllati e thesauri, consente una descrizione agevolata e accurata delle risorse. Per le finalità specifiche del presente progetto di ricerca sono stati presi a riferimento, oltre al già citato Metadata Standards Crosswalk del Getty Institute, i *crosswalk* specifici VRA Core/DC definiti dalla American Museum of Natural History¹⁴⁸.

4. Analisi delle immagini e riconoscimento automatico di immagini statiche dinamiche e audiovideo: metodologie e strumenti

Le tecnologie di intelligenza artificiale (AI) hanno contribuito, negli ultimi decenni, allo sviluppo di strumenti di analisi dei documenti per estrarne efficacemente frammenti di informazione. I maggiori progressi si registrano nell'analisi e riconoscimento di forme e oggetti nelle immagini. Naturalmente tale approccio, se bene implementato potrebbe risultare particolarmente utile alla gestione di documenti di archivio e più in generale certamente alla valorizzazione e diffusione del patrimonio culturale. Gli strumenti realizzati sono vari e, perlopiù, si basano su modelli statistici in grado di indagare il contenuto delle immagini o di documenti multimediali, da cui estrarre elementi informativi: classi di metadati, entità testuali, oggetti visivi, argomenti etc.

Il Machine Learning (ML) – Apprendimento automatizzato – è un branca dell'Intelligenza Artificiale dedicata alla costruzione di algoritmi in grado di replicare l'apprendimento umano¹⁴⁹. Gli algoritmi di ML sono implementati per elaborare diverse attività quali, ad esempio, classificazione, clusterizzazione, regressione, individuazione di anomalie etc.

All'interno del machine learning si distinguono tre principali tipologie di apprendimento: supervisionato, non supervisionato e per rinforzo.

Apprendimento supervisionato

L'apprendimento supervisionato è una sottocategoria di machine learning e AI. È definito dal suo utilizzo di set di dati etichettati per addestrare algoritmi che classificano i dati o prevedono accuratamente i risultati. Man mano che i dati di input vengono immessi nel modello, esso adegua i suoi coefficienti fino a quando il fitting del modello non sarà stato eseguito in modo appropriato, il che avviene come parte del processo di convalida incrociata.

¹⁴⁸ American Museum of Natural History, "Dublin Core to VRA Core Mapping," ultimo accesso 23 Gennaio 2024, <https://sites.google.com/site/dctovrc/home>.

¹⁴⁹ Shinde, Pramila P., Shah, Seema. "A Review of Machine Learning and Deep Learning Applications". *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*, 2018. DOI: 10.1109/ICCUBEA.2018.8697857.

Apprendimento non supervisionato

l'apprendimento non supervisionato avviene quando i dati di input non sono etichettati o non hanno un corrispondente valore di output e l'algoritmo deve scoprire eventuali relazioni esistenti. Il suo utilizzo è adatto a cercare modelli nascosti nei dati che sfuggono all'osservazione sia perché oscurati da altre informazioni sia perché la quantità di dati è talmente grande da non poter essere osservata facilmente senza un ausilio computazionale. Senza supervisione il modello non può essere addestrato su un insieme di dati preparati con il corretto output corrispondente ma deve individuare autonomamente le differenze o le similitudini fra gli input identificandone le caratteristiche principali; non dovendo effettuare le fasi di addestramento e di validazione la preparazione del dataset di input richiede meno sforzi.

Apprendimento per rinforzo

Infine, l'apprendimento di rinforzo – *reinforcement learning* – mira a creare un sistema in cui l'algoritmo è in grado di migliorarsi confrontandosi con l'ambiente in cui è posto. In particolare, il sistema prevede l'uso di premialità e penalizzazioni per agevolare il processo di addestramento.

Il reinforcement learning si utilizza nei casi in cui è necessario raggiungere un obiettivo in un ambiente incerto, quando non è possibile prevedere tutte le variabili.

Deep Learning

Il Deep Learning (DL) costituisce un sotto campo del Machine Learning in cui l'utilizzo di reti neurali (*neural networks*) rende possibile implementare attività che emulano il funzionamento del cervello umano e le sinapsi dei neuroni. Tali reti sono composte da nodi (*nodes*), nei quali avviene il processamento dei diversi dati in input, organizzati su più strati (*layers*) distinti generalmente in:

- **Input layer:** il livello dei dati in ingresso da processare.
- **Hidden layers** (strati o livelli nascosti): costituiscono il livello in cui avviene l'elaborazione dei dati nelle architetture delle reti neurali.
- **Output layer:** ultimo livello dell'architettura, all'interno del quale i dati provenienti dagli *hidden layer* vengono processati per restituire il risultato finale pianificato.

Si possono distinguere due tipologie principali reti neurali: *Recurrent Neural Network(s)* (RNNs) e *Convolutional Neural Network(s)* (CNNs). Le prime, tradotte letteralmente come reti neurali ricorrenti, sono una classe di architetture di reti neurali per l'elaborazione di sequenze di dati di vario genere. In particolare, questo tipo di reti utilizza i valori di output di uno strato superiore come dati di input di un livello inferiore in modo tale che ad ogni strato più profondo la rete acquisisce la capacità di riconoscere caratteristiche più complesse, ricombinando i dati dei nodi precedenti. Tali architetture

sono largamente applicate ad es. in attività di riconoscimento automatico della scrittura (*Handwritten Text Recognition*) di cui si parlerà più avanti.

Le reti neurali convoluzionali costituiscono l'architettura di base nel campo della Computer Vision il cui scopo è quello di ricavare informazioni dai vari input visivi. Le CNNs sono infatti in grado di processare le immagini statiche e dinamiche, al fine di classificarle, riconoscerle le forme, estrarne il contenuto informativo.

Immagini statiche e video

Applicate alle immagini, siano esse statiche o dinamiche, le reti neurali convolutive consentono l'analisi e l'estrazione di contenuti informativi. Si può parlare, in questo caso, di riconoscimento poiché l'obiettivo dei modelli sviluppati tramite questa architettura è, per l'appunto, quello di riconoscere e classificare oggetti/elementi presenti nelle immagini e nei video. Comunemente si distinguono tre principali operazioni effettuate sulle immagini statiche e dinamiche tramite le reti convolutive:

- *Image/video classification.*
- *Object detection*
- *Object segmentation*

In questo paragrafo si illustreranno queste attività sotto un aspetto metodologico e strutturale individuando alcuni tra i principali modelli sviluppati¹⁵⁰.

Image Classification

La classificazione è un processo logico ben sviluppato anche nelle tecniche di apprendimento automatico supervisionato. In questo caso è necessario che l'oggetto da classificare sia dotato di caratteristiche, "*features*", sulle quali addestrare le reti neurali per calcolare la probabilità di appartenenza ad una specifica classe. Tradizionalmente, esistono due principali modalità di classificazione:

1. **Binaria**, quando la scelta deve essere effettuata tra due classi (es. 0-1, bianco-nero etc.)
2. **Multiclasse**, in tal caso la scelta è operata tra tre o più classi (tipologie di oggetti, animali, piante etc.).

In base al tipo di classificazione è possibile applicare tecniche di Machine Learning con specifici algoritmi. Tra questi ricordiamo:

¹⁵⁰ Per un dettaglio tecnico dei processi coinvolti e della costruzione algoritmica si rinvia a letture specifiche come Weidman, Seth. *Deep Learning from Scratch. Building with from First Principles.* O'Reilly, 2019.

- **Linear regression**, un modello lineare particolarmente efficace per la classificazione binaria. Si basa «sull'assunzione che esista una relazione lineare tra le variabili in input e le probabilità di appartenenza ad una classe»¹⁵¹;
- **Support Vector Machine**, è un algoritmo impiegato nella classificazione multiclasse per determinare il margine del classificatore, ossia la linea di demarcazione nell'iperpiano il più distante possibile tra le istanze delle classi presenti. È un algoritmo di alta precisione, sia in problemi di classificazione lineare che non lineare, che, di contro, può richiedere una elevata potenza computazionale se applicato su grandi dataset¹⁵². Questo algoritmo è alla base delle tecniche di *Sentiment Analysis*, *Spam detection* e *Text category assignment*.
- **Decision Tree**, è un algoritmo di classificazione non lineare, ovvero è in grado di catturare le relazioni non lineari tra le caratteristiche del dataset e le etichette delle classi. L'algoritmo procede verso la classificazione dell'istanza mediante la costruzione di un modello decisionale a forma di albero guidato dalle caratteristiche dei dati. Il processo decisionale prosegue sino a quando non viene raggiunto un punto di arresto e l'istanza viene assegnata ad una determinata classe¹⁵³.

Nel caso specifico della classificazione applicata alle immagini, l'elaborazione diviene più complessa a causa del numero elevato di input da processare. Le reti neurali convolutive (CNNs) consentono di gestire questi processi con minori difficoltà grazie alla possibilità di elaborare, contemporaneamente, diversi input. Sono ampiamente adottate *Computer vision* con l'obiettivo di simulare il sistema visivo umano in applicazioni quali, ad esempio: riconoscimento di oggetti e testi da immagine, sviluppo dei sistemi a guida autonoma, lettura di immagini mediche etc.

A partire dalla struttura di base delle reti neurali (*input layer*, *hidden layer*, *output layer*), le CNNs sono contraddistinte da:

- *Convolutional layer*: l'operazione di convoluzione ha la finalità di definire una “*features map*”, ovvero una mappa delle caratteristiche, che individui tra i pixel dell'immagine in input le caratteristiche essenziali per classificarla. La componente principale della convoluzione è il *kernel* o *filtro*, la matrice di pesi che viene sovrapposta ai pixel dell'immagine per l'estrazione delle *features*¹⁵⁴. In ogni sovrapposizione i valori dei pixel vengono moltiplicati e sommati con quelli del *kernel* per ottenere in output la mappa delle caratteristiche che

¹⁵¹ Cristian Tesconi, *Intelligenza Artificiale* (2023), p. 57. ISBN: 979-8862365245.

¹⁵² Cristian Tesconi, *Intelligenza Artificiale cit.*, p. 75.

¹⁵³ Cristian Tesconi, *Intelligenza Artificiale cit.*, p. 70.

¹⁵⁴ In particolare, il kernel viene applicato procedendo per sovrapposizioni successive ai dati di input, che costituiscono porzioni dell'immagine.

evidenzia i punti più rilevanti dell'immagine. La convoluzione costituisce il *layer* centrale all'interno dell'architettura delle CNNs da cui il nome.

- *Pooling layer*: agisce sull'output del livello precedente per ridurre il numero di parametri che la rete dovrà continuare ad analizzare. Come per la convoluzione, anche in questo caso viene applicato un filtro a cui non è associato alcun peso. A seconda del tipo di pooling utilizzato, il filtro agisce in modo da individuare il pixel con il valore massimo o con il valore medio per ogni parte analizzata, che comporrà l'output di questa operazione. Uno dei vantaggi del pooling consiste nel diminuire il numero di caratteristiche da processare con la conseguente riduzione di potenza computazionale richiesta per l'elaborazione dell'immagine. All'interno delle CNNs non vi è limite di costruzione dei blocchi, pertanto, possono essere definiti più strati di convoluzione o di *pooling*, al termine dei quali viene configurato lo strato di classificazione da cui ottenere il risultato finale.
- *Fully Connected Layer*: La parte finale di ogni rete, che deve combinare tutte le caratteristiche estratte negli strati precedenti per produrre il risultato finale. Tali strati possono essere adottati in numero variabile, in base alla topologia di rete. Gli FCL hanno la caratteristica di connettere ogni nodo dell'output a un nodo del livello precedente.
- *Loss function*. Nota anche come *error function*, è una funzione che definisce l'accuratezza del modello, delineandone la capacità predittiva con la misurazione della discrepanza tra i valori predetti e i valori reali di riferimento (groundtruth). È possibile utilizzare diverse loss function a seconda della classificazione, comunemente si adotta la *Binary cross-entropy* per la classificazione binaria e la *Categorical cross-entropy* per la classificazione multiclasse.

A titolo esemplificativo, nella Figura 13 si riporta l'esempio di una immagine in scala di grigi (28x28x1) processata da una generica architettura CNN. La struttura si compone di un iniziale strato convoluzionale con un *kernel* di 5x5 e una *feature maps* in output delle dimensioni di 32 seguito uno strato di *pooling* 2x2. Successivamente, viene definito un secondo strato convoluzionale di dimensioni maggiori, 5x5x64, e un nuovo strato di *pooling* 2x2. L'output risultante da queste operazioni rappresenta un nuovo input in uno strato di *Fully connected layer* (FC). L'ultima fase di processamento avviene a livello di *Fully connected layer* all'interno del quale prende avvio la funzione di attivazione *softmax* per ottenere il risultato finale della probabilità di appartenenza dell'immagine processata ad una determinata classe.

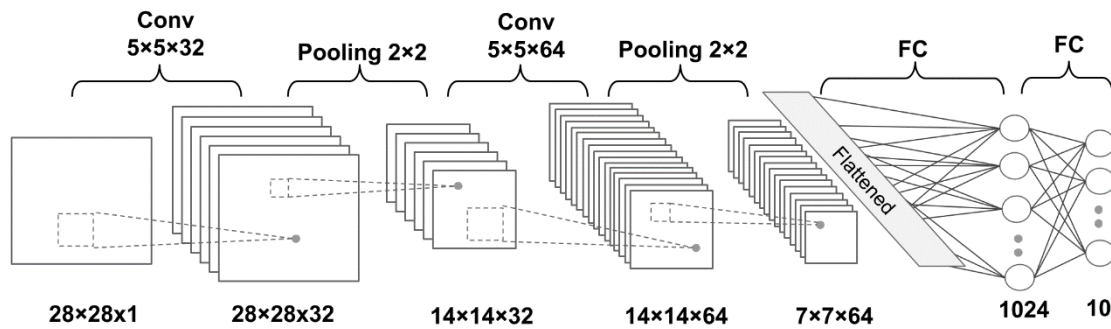


Figura 13: architettura di rete neurale convoluzionale¹⁵⁵.

Nel corso degli anni sono state implementate un cospicuo numero di architetture che hanno progressivamente migliorato l'accuratezza nelle operazioni di riconoscimento e classificazione. In particolare, lo sviluppo è stato possibile mediante le *challenges* promosse a livello internazionale per testare le capacità dei modelli prodotti. Le competizioni hanno lo scopo di stabilire dei benchmark di prestazione e costituire una base di confronto per la valutazione dei sistemi proposti. Tra le principali competizioni promosse si cita:

1. *Pattern Analysis, Statistical Modelling and Computational Learning Visual Object Classes (PASCAL VOC)*¹⁵⁶. Una delle prime competizioni, svolta dal 2005 al 2012, per il confronto tra i modelli per il riconoscimento degli oggetti. Nell'ultimo anno della competizione, il dataset si compone di circa 12.000 immagini e venti classi di oggetti.
2. *ImageNet Large Scale Visual Recognition Challenge (ILSVRC)*¹⁵⁷. Svolta dal 2010 sino al 2017, la competizione ha avuto per oggetto il confronto tra modelli per la classificazione di immagini e il riconoscimento di oggetti. Il dataset è tra i più estesi al mondo e contiene circa 15 milioni di immagini e 1.000 categorie di oggetti.
3. *Common Objects in Context (COCO)*¹⁵⁸. La competizione si è svolta dal 2014 sino al 2020 e ha avuto per oggetto il confronto dei modelli per il riconoscimento degli oggetti e altre attività di *Computer Vision*. Il dataset si compone di circa 300.000 immagini e 80 categorie di oggetti.

Attraverso i risultati raggiunti tramite le competizioni, alcuni modelli hanno costituito per diverso tempo un punto di riferimento per l'implementazione di nuove soluzioni. Tra questi, si cita a titolo esemplificativo, alcuni dei modelli che hanno caratterizzato il percorso evolutivo di tali tecnologie.

¹⁵⁵ Roshka, Sebastian, and Vahid Mirjalili. *Python Machine Learning. Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow 2*. Ed. Terza. Birmingham. Pack Publishing. 2019, p. 542.
¹⁵⁶ Mark Everingham et al., "The PASCAL Visual Object Classes Challenge: A Retrospective," *International Journal of Computer Vision* 111, n. 1 (2015). DOI 10.1007/s11263-014-0733-5. <http://host.robots.ox.ac.uk/pascal/VOC/>
¹⁵⁷ Olga Russakovsky et al., "ImageNet Large Scale Visual Recognition Challenge," *IJCV*, (2015). <https://doi.org/10.48550/arXiv.1409.0575>.
¹⁵⁸ Tsung-Yi Lin et al., "Microsoft COCO: Common Objects in Context," (2015). <https://arxiv.org/abs/1405.0312>. <https://cocodataset.org/#home>.

AlexNet¹⁵⁹ ha rappresentato una vera innovazione nel campo delle CNN raggiungendo un'accuratezza¹⁶⁰ dell'84,6%¹⁶¹. La caratteristica principale di AlexNet è la sua profondità, composta da otto strati neurali, di cui cinque sono strati convoluzionali e tre sono completamente connessi. L'architettura, progettata per lavorare con immagini di grandi dimensioni, applica una tecnica di regolarizzazione per non incorrere in problemi di *overfitting*¹⁶².

Le reti VGG¹⁶³, *Visual Graphics Group*, hanno apportato una maggiore profondità nella definizione dell'architettura arrivando a contenere 16 livelli nella versione VGG-16 (13 strati convoluzionali e 3 FCL) e 19 nella versione VGG-19 (16 strati convoluzionali e 3 FCL). L'articolazione profonda dell'architettura permette l'elaborazione di un maggior numero di caratteristiche, arrivando a raggiungere un'accuratezza, sul dataset *ImageNet*, pari al 92,3%¹⁶⁴.

Allo stato attuale, GoogleNet¹⁶⁵ è tra i modelli che hanno raggiunto le migliori prestazioni con un top-5 accuracy del 93,3%¹⁶⁶. Uno dei maggiori vantaggi prodotti da questa rete è l'ottimizzazione delle risorse del computer attraverso l'apporto del modulo "*inception*" che processano le operazioni di convoluzione senza un aumento esponenziale del numero di parametri.

Uno degli aspetti negativi connessi all'implementazione di soluzioni con un elevato numero di strati di processo è connesso con la "perdita del gradiente" (*Vanishing gradient*) che può portare ad una interruzione dell'addestramento. Una prima soluzione a tale problema si ha con l'implementazione di ResNet¹⁶⁷, *Residual Network*, che sfrutta blocchi "residuali" per ottimizzare le connessioni tra gli

¹⁵⁹ Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Advances in Neural Information Processing Systems 25 (NIPS 2012)*, (2012). https://papers.nips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf.

¹⁶⁰ L'accuratezza, espressa nelle misure di Top-1 e Top-5, è stata valutata tenendo conto della complessità computazionale, calcolata in *Floating-point operations* (FLOPs).

¹⁶¹ Anamika Dhillon, and Gyanendra K. Verma, "Convolutional neural network: a review of models, methodologies and applications to object detection," *Progress in Artificial Intelligence* 9, (2020). DOI: 10.1007/s13748-019-00203-0.

¹⁶² L'*overfitting* si verifica nel momento in cui il modello risulta eccessivamente adattato al dataset di addestramento e non è in grado di processare correttamente i nuovi input. Al contrario, l'*underfitting* si verifica nel momento in cui il modello non è in grado di comprendere la relazione tra dati in input e in output comportando errori anche in fase di addestramento.

¹⁶³ Karen Simonyan, and Andrew Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," (2015). <https://arxiv.org/abs/1409.1556>.

¹⁶⁴ Genevieve Sapijaszko, and Wasfy B. Mikhael, "An Overview of Recent Convolutional Neural Network Algorithms for Image Recognition," *2018 IEEE 61st International Midwest Symposium on Circuits and Systems (MWSCAS)*, (2018), p. 2. DOI: 10.1109/MWSCAS.2018.8623911.

¹⁶⁵ Christian Szegedy et al., "Going deeper with convolutions," (2014). <https://arxiv.org/abs/1409.4842>.

¹⁶⁶ Anamika Dhillon, and Gyanendra K. Verma, "Convolutional neural network: a review of models, methodologies and applications to object detection," *cit.*, p. 92.

¹⁶⁷ Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Deep Residual Learning for Image Recognition," (2015). <https://arxiv.org/abs/1512.03385>.

strati e, così facendo, risultare più efficiente da un punto di vista computazionale ottenendo i medesimi livelli di accuratezza delle reti citate in precedenza¹⁶⁸.

Per un corretto funzionamento, queste tipologie di architetture necessitano di grandi quantità di addestramento che, talvolta, non è possibile ottenere. In questi casi si procede ad applicare tecniche di *data augmentation* per incrementare il volume dei dati a disposizione tramite la duplicazione e modifica delle immagini già presenti al fine di ottenere un numero maggiore di esempi. Modificando i duplicati delle immagini, ad esempio con la rotazione, saturazione, contrasto etc., gli algoritmi rileveranno quelle modifiche come esempi aggiuntivi rispetto agli originali.

Il task di classificazione costituisce la base dei processi di computer vision non solo per l'analisi delle immagini statiche ma anche per quelle dinamiche. La classificazione può essere applicata alle immagini dinamiche per riconoscere ad es. le attività rappresentate nel video. In questo specifico caso, vengono applicati algoritmi in grado di identificare la classe di attività a cui associare il video sulla base di un approccio probabilistico. Al fine di limitare il margine di errore viene elaborata la relazione spazio-temporale tra fotogrammi prendendo in considerazione, nella classificazione di ciascun fermo immagine, n frame successivi e precedenti con l'obiettivo di contestualizzare maggiormente il fotogramma in esame e migliorare la precisione della classificazione.

Object recognition

L'attività di object recognition è finalizzata all'individuazione e al riconoscimento di oggetti/elementi all'interno di immagini statiche e dinamiche prevedendo una fase di riconoscimento attraverso la segmentazione ed una successiva di associazione alle classi. Gli algoritmi di *object detection* sono in grado di affinare il processo di classificazione attraverso l'addestramento. In particolare, il termine object detection viene usato per indicare il processo volto all'individuazione di più istanze di un oggetto, combinando le funzionalità di object classification con quelle di object localization (individuazione della regione che contiene l'oggetto). Il risultato finale è l'immagine con l'evidenziazione dell'oggetto tramite un riquadro, propriamente denominato *bounding box* (Figura 14).

¹⁶⁸ Musab Coşkun, Özal Yildirim, Ayşegül Uçar, and Yakup Demir, "An Overview of Popular Deep Learning Methods," *European Journal of Technique (EJT)* 7, n. 2 (2017), p. 5. <https://dergipark.org.tr/en/download/article-file/437659>.

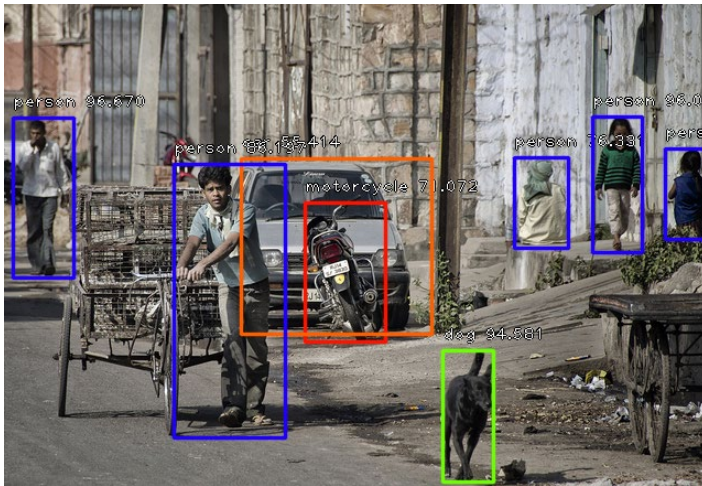


Figura 14: esempio di output di object detection¹⁶⁹.

Nel campo dell'object detection, possono essere utilizzati molteplici approcci di CNN per identificare le coordinate di un oggetto e la sua classe, alcuni dei più diffusi sono: *Region Proposal Based*, *Regression/Classification Based*¹⁷⁰.

Il *Region Proposal Based*, comune alla maggior parte delle architetture, si sostanzia in un processo distinto in due fasi che include l'individuazione di "regioni" per delineare elementi all'interno delle immagini. L'idea alla base della Region proposal Network è quella di utilizzare la mappa delle features ottenuta dai layer convoluzionali per generare delle "aree proposte", che abbiano la probabilità di contenere degli oggetti. L'applicazione delle regioni consente, inoltre, un notevole risparmio computazionale. Girshick et al.¹⁷¹ sono stati i principali artefici di questa tecnica implementando l'architettura *Region-Based CNN* (R-CNN) che vede l'impiego di un algoritmo di ricerca selettiva¹⁷² per la generazione di 2000 regioni per immagine. L'algoritmo, dopo aver segmentato l'immagine, sfrutta l'intensità dei pixel per individuare zone simili fra loro. Individuata ciascuna zona, il R-CNN prevede l'applicazione di reti convolutive per l'estrazione delle *features* e la classificazione finale di ciascun segmento tramite un algoritmo di *Support Vector Machine* (Figura 15). Questa architettura è stata addestrata con un approccio supervisionato sul dataset ILSVRC¹⁷³ e ha costituito per diverso tempo la base da cui partire per l'implementazione di nuove architetture.

¹⁶⁹ Olafenwa Moses. "ImageAI: Object Detection." <https://github.com/OlafenwaMoses/ImageAI/blob/master/data-images/image3new.jpg>

¹⁷⁰ Zhong-Qiu, Zhao, Xu Shou-Tao, and Wu Xindong. Object Detection with Deep Learning: A Review. 2019, p. 3. doi:/10.48550/arXiv.1807.05511.

¹⁷¹ Girshick, Ross, Jeff Donahue, Trevor Darrell, and Jitendra Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. *Computer Vision and Pattern Recognition*. 2014. <https://arxiv.org/abs/1311.2524>.

¹⁷² Uijlings, Jasper, K. E. Van De Sande, T. Gevers, and A. W. Smeulders. Selective search for object recognition. *International Journal of Computer Vision* 104, n. 2. 2013. pp. 154–171.

¹⁷³ ImageNet. "Large Scale Visual Recognition Challenge." <https://www.image-net.org/challenges/LSVRC/>

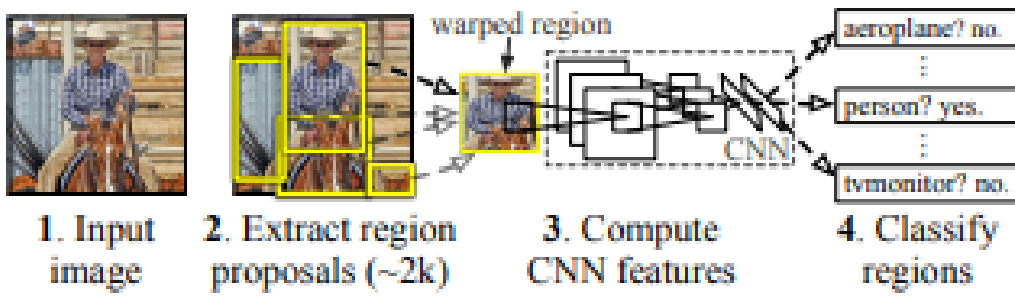


Figura 15: struttura del R-CNN.¹⁷⁴

Uno dei principali limiti derivanti da questa architettura è il costo, in termini di memoria e potenza computazionale, richiesto durante la fase di addestramento oltre alla ridondanza nella generazione delle regioni che contribuiscono ad elevare i tempi di calcolo. Per superare parte di questi svantaggi è stata realizzata una versione più performante della R-CNN, la Fast R-CNN.

L'eventuale applicazione del *Region of Interest pooling layer* (RoI)¹⁷⁵ (Figura 16), all'interno di architetture *Fast R-CNN*, permette di ottenere regioni uniformi che comprimono i tempi di calcolo, poiché la generazione delle regioni avviene successivamente allo strato convoluzionale e alla creazione delle *feature maps*, .

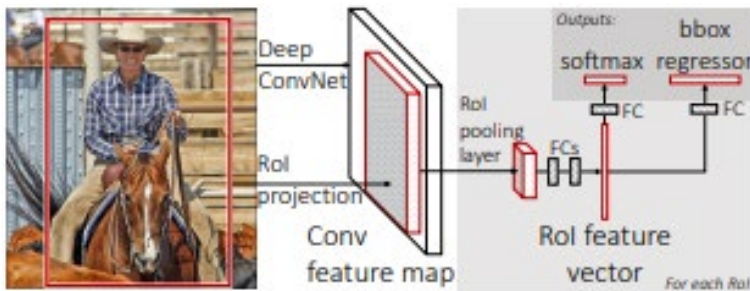


Figura 16: Struttura del Fast R-CNN¹⁷⁶

Fast R-CNN impiega l'algoritmo di ricerca selettiva per ottenere le proposte di regioni passando dai 50 secondi circa di *R-CNN* a 2 secondi circa. Inoltre Ren et al.¹⁷⁷ hanno proposto una tecnologia *Faster R-CNN* che, risulta essere ancora più veloce rispetto alle precedenti *Fast R-CNN* e *R-CNN*. L'idea di fondo di tale architettura è la sostituzione dell'algoritmo di ricerca selettiva esterno con un interno *Region Proposal Network* (RPN) in grado di apprendere e successivamente generare le proposte di regioni. Il RPN procede dalle *feature maps* per predire i riquadri contenenti entità o sfondo da restituire come input per lo strato di *Region of Interest pooling*. L'introduzione di questa

¹⁷⁴ Girshick, Ross, Jeff Donahue, Trevor Darrell, and Jitendra Malik. *cit.* p.1.

¹⁷⁵ Girshick Ross. Fast r-cnn. 2015. arXiv:1504.08083.

¹⁷⁶ Girshick Ross. Fast r-cnn. 2015. arXiv:1504.08083. p.2.

¹⁷⁷ Ren, Shaoqing, Kaiming He, Ross Girshick, Jian Sun. Faster r-cnn: Towards realtime object detection with region proposal networks. 2015. <https://arxiv.org/abs/1506.01497>. pp. 91–99.

“sottorete” all’interno dell’architettura permette di diminuire notevolmente i tempi di elaborazione dai 2 secondi previsti per ciascuna immagine con *Fast R-CNN* ai 0.2 secondi.

Le architetture basate sull’individuazione delle regioni si compongono di diverse fasi, addestrate separatamente, tra cui rientrano il *region proposal*, l’estrazione delle caratteristiche, la classificazione e la definizione dei *bounding boxes*. Nonostante i progressi effettuati e i risultati raggiunti, queste architetture necessitano di un’attenzione maggiore nella gestione delle varie componenti.

L’approccio basato sulla classificazione o regressione, noto anche come *single shot detectors*, consiste nel mappare direttamente i pixel dell’immagine evidenziandone i riquadri delimitatori e calcolandone, contestualmente, la probabilità di classificazione delle istanze. Si tratta di modelli che cercano di completare il processo di riconoscimento in un passaggio unico, utilizzando un’unica rete. Sulla base delle performance ottenute fino ad oggi, la principale architettura, che realizza un approccio di questo tipo è *You Only Look Once* (YOLO) di cui sono stati implementati diversi modelli¹⁷⁸ altamente efficienti. Proposta originariamente nel 2015 da Redmon et al.¹⁷⁹, YOLO ha ottenuto grande successo in particolare nel riconoscimento di oggetti in tempo reale per la particolare velocità di elaborazione di cui è capace. Nella versione iniziale, il punto iniziale è la suddivisione dell’immagine in celle $N \times N$ per ognuna delle quali viene predetta la classe di ogni singolo oggetto presente. Ogni cella della griglia è responsabile per la predizione del *bounding box* intorno all’oggetto riconosciuto con la relativa classe (la più quotata). La formula

$$Pr(Object) * IOU * Pr(Class_i | Object)$$

calcola la probabilità che l’oggetto esista all’interno del riquadro. Nella formula il modello, per ogni *bounding box*, moltiplica le probabilità condizionate delle classi (della cella) con il confidence score della *bounding box*, ottenendo un confidence score specifico di ogni classe per ogni *bounding box*. IOU= Intersection Over Union è una metrica utilizzata nei sistemi di object detection per determinare lo scostamento tra le *bounding box* previste dal modello e quelle reali.

L’attuale versione prodotta da Ultralytics¹⁸⁰, YOLO v8, costituisce il modello di riferimento della serie YOLO risultando più veloce e più accurato rispetto ai precedenti¹⁸¹. L’architettura è disponibile in cinque diverse versioni differenziate per complessità della struttura partendo dalla più semplice e veloce, YOLOv8n, fino alla più estesa e accurata YOLOv8x.

¹⁷⁸ Ad oggi si contano ben otto versioni e diverse varianti: YOLO v1, YOLO v2, YOLO v3, YOLO v4, YOLO v5, YOLO v6, YOLO v7 e YOLO v8.

¹⁷⁹ Redmon, Joseph, Santosh Divvala, Ross Girshick, and Ali Farhadi. You Only Look Once: Unified, Real-Time Object Detection. 2016. arxiv.org/pdf/1506.02640.

¹⁸⁰ Jocher, Glenn. “Ultralytics.” <https://github.com/ultralytics/ultralytics>. Ultimo accesso Maggio 2023.

¹⁸¹ «Ultralytics YOLOv8 is the highest performing YOLO model ever released, setting new standards in SOTA real-time detection and segmentation». Jocher, Glenn. “Ultralytics.” <https://ultralytics.com/yolov8>. Ultimo accesso Maggio 2023.

Lo stato dell'arte sulle reti convolutive, in particolare per l'*object detection*, è in continua evoluzione generando nuovi e sempre più copiosi modelli, orientati a sviluppare specifici aspetti di computer vision, quali ad esempio:

- **3D Object Detection**, è il task fondamentale per la guida automatizzata e la robotica e prevede l'identificazione di oggetti all'interno di un ambiente tridimensionale. Per questo processo vengono utilizzati comunemente dataset quali KITTI, Waymo Open Dataset, nuScenes e ScanNet.
- **Real-Time Object Detection**, è il task volto all'individuazione di oggetti nei video in diretta capace di combinare tecniche di tracciamento e *object detection*. Come anticipato in precedenza, la famiglia YOLO è l'architettura maggiormente applicata per il riconoscimento degli oggetti nei video in tempo reale.
- **Salient Object Detection**, è un'attività di attenzione visiva, basata su un meccanismo di percezione che permette di localizzare gli oggetti o le regioni di maggiore attrazione all'interno di una immagine. In base alla specificità della tipologia di immagine, il task assume una propria declinazione come *Video Salient Object Detection* nel caso di immagini dinamiche o *RGB Salient Object Detection* con immagini a colori.

Object segmentation

Altra tecnica rilevante per la computer vision è l'*object segmentation*, che mira a suddividere l'immagine in parti più elementari fino a individuarne forma e contorno.

L'esempio nella Figura 17 mostra la differenza tra il processo di riconoscimento (a sinistra), *instance segmentation* (al centro) e *whole-scene semantic segmentation* (a destra).

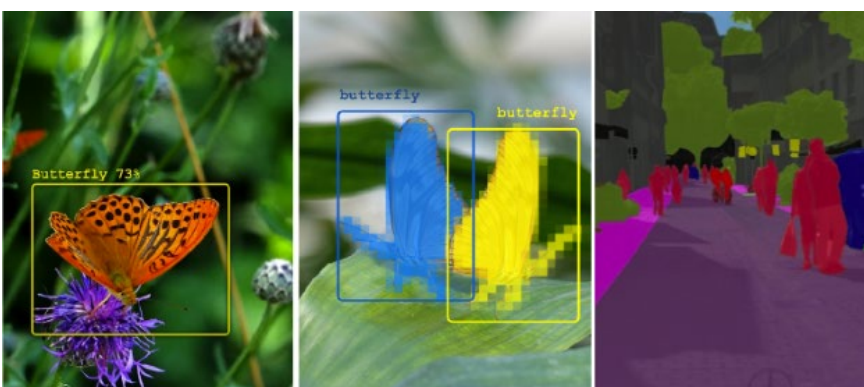


Figura 17: Esempio di risultato di *object detection* (sinistra), *instance segmentation* (centro) e *whole-scene semantic segmentation* (destra)¹⁸².

Rispetto al processo di riconoscimento che è stato trattato precedentemente, la segmentazione è un'operazione aggiuntiva che consiste nel marcare i singoli pixel che compongono l'entità

¹⁸² Lakshmanan, Valliappa, Martin Görner, and Ryan Gillard. *Practical Machine Learning for Computer Vision*. O'Reilly Media Inc. 2021. p. 131.

individuata e localizzata. Come risulta dalla Figura 17, la segmentazione può prevedere da un lato *instance segmentation* che focalizza l'attenzione esclusivamente sugli oggetti presenti nell'immagine oppure *whole-scene semantic segmentation* che analizza e processa l'intera immagine segmentando tutto il contenuto presente.

Nel primo caso, si utilizza l'approccio *two-stage* impiegando quindi reti per la generazione delle regioni per individuare con maggiore precisione gli oggetti da cui partire per la fase di segmentazione. L'architettura *Mask R-CNN*¹⁸³ costituisce un punto di riferimento per l'elaborazione di questa attività ed è stata implementata come un'estensione di *Faster R-CNN* a cui aggiunge, parallelamente alla predizione della classe, la generazione di una maschera in bianco e nero della *feature map* per individuare il contorno dell'oggetto. Di fatto, questa maschera costituisce il nucleo dell'architettura tramite la quale individuare i singoli pixel che compongono l'oggetto.

Nel caso della "segmentazione semantica" l'obiettivo è ben più ampio e volto a classificare ciascun pixel dell'immagine. Un'architettura largamente impiegata per questa attività è U-Net¹⁸⁴ implementata originariamente nel settore biomedico per l'individuazione delle entità nelle immagini da microscopio. Come mostrato nella Figura 18, la struttura si compone di un codificatore (lato sinistro) connesso con un decodificatore (lato destro) tramite un ponte di collegamento. La struttura così definita consente di effettuare una prima operazione di ridimensionamento dell'immagine ed estrazione delle caratteristiche a cui fa seguito il progressivo ingrandimento dell'immagine e concatenazione delle informazioni. Il passaggio dalla fase di codifica a quella di decodifica consente l'estrazione di informazioni a diversi livelli di granularità e di raffinare la segmentazione degli oggetti. In particolare, l'*up-sampling* elaborato nel decodificatore permette di mantenere le informazioni contenute nelle aree di dimensioni notevolmente ridotte.

¹⁸³ Kaiming He, Georgia Gkioxari, Piotr Dollar, and Ross Girshick. Mask R-CNN. 2018. arXiv:1703.06870.

¹⁸⁴ Ronneberger Olaf, Philipp Fischer, and Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. 2015. arXiv:1505.04597v1.

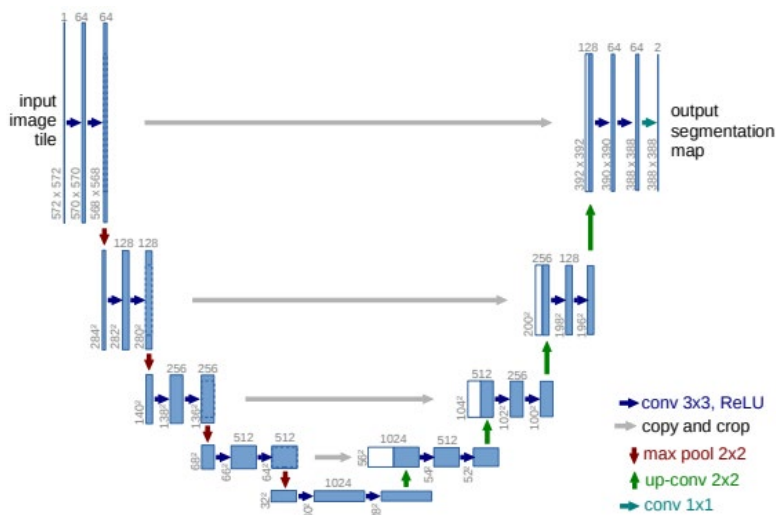


Figura 18: struttura di U-Net¹⁸⁵.

Reti Neurali e riconoscimento di testi manoscritti digitalizzati

Negli ultimi anni si è assistito ad una progressiva e crescente digitalizzazione di testi manoscritti.

Il riconoscimento della scrittura a mano in immagine di testi e documenti rappresenta un tema complesso, con luci e ombre e ancora lontano da soluzioni soddisfacenti definitive.

L'attività di elaborazione della scrittura a mano da immagini, Intelligent Character Recognition (ICR), si distingue dall'OCR (Optical Character Recognition) che ha per oggetto la lettura dei caratteri a macchina. Rispetto a quest'ultimo, la sfida posta in essere nell'ICR è la capacità di riconoscere lettere e simboli che presentano forme di scrittura differenti e soggette a diversi fattori (supporto scrittorio, stato emotivo dello scrivente, etc.). Si tratta di un'attività per la quale le reti neurali sono impiegate per processare ingenti quantità di dati a seconda che si tratti di testi strutturati, es. la compilazione di un modulo, o non strutturati, es. testo libero o appunti o manoscritti storici. Unitamente alla varietà dello stile di scrittura, la sintassi grammaticale della lingua costituisce un fattore decisivo per la corretta riuscita del riconoscimento.

Tradizionalmente, i sistemi di ICR basavano il riconoscimento della scrittura classificando le parole estrapolate in base ai lessici di riferimento (Lexicon-based), dimostrando una elevata accuratezza legata indissolubilmente alle risorse terminologiche specifiche di un determinato dominio (lessici, thesauri, vocabolari ecc.). Nell'esecuzione del processo, i sistemi effettuano una preliminare segmentazione del file finalizzata a individuare le righe di testo da cui partire per il riconoscimento di simboli e lettere. A differenza dell'OCR, in cui lo spazio delimitante i segni grafici è facilmente individuabile, nell'ICR tale attività è più complessa comportando, di conseguenza, il dispendio di maggiore potenza di calcolo per processare intere righe di testo.

¹⁸⁵ Ronneberger Olaf, Philipp Fischer, and Thomas Brox. *cit.* p.2.

Nel corso degli anni, la ricerca è stata orientata verso l'applicazione, anche congiunta, di Recurrent Neural Networks (RNNs) e Convolutional Neural Networks (CNNs) che consentono di rendere il riconoscimento dei caratteri e dei simboli sempre meno dipendente dall'uso di lessici o vocabolari di riferimento, i quali possono entrare nel processamento solo alla fine per affinare la qualità semantica di quanto riconosciuto. L'impiego delle RNNs come Long-Short Term Memory (LSTM), grazie alla capacità di mantenere una memoria a breve e lungo termine, permette una migliore comprensione del contesto del testo rispetto a modelli come Hidden Markov Model (HMM) che dimostrano poca affidabilità nel gestire lunghe sequenze di parole interdipendenti tra loro. La combinazione con le reti neurali convoluzionali consente di ottimizzare ulteriormente il processo mediante una preliminare attività finalizzata a selezionare, e dunque elaborare, le sole zone in cui è contenuto del testo. Un ulteriore fattore di cui si è tenuto conto ha riguardato la possibilità di diminuire i tempi generali di calcolo riconoscendo termini da segni grafici simili tra loro. Sulla base di tali considerazioni, sono state sviluppate metodologie per estrarre n-grammi ed elaborare un riconoscimento a livello di termini.

Di recente, sono stati sviluppati sistemi basati interamente sulle reti neurali convoluzionali per i quali, la presenza di lessici o vocabolari di riferimento non costituisce un elemento necessario ma opzionale per, eventualmente, incrementare l'accuratezza del riconoscimento.

A tal proposito, il sistema descritto da Ptucha et al. (2019) prevede l'implementazione di una rete Lexicon CNN con la funzione di predire le corrispondenze delle parole individuate con il lessico di riferimento. Nel caso specifico, il confronto con il lessico è funzionale per l'ottimizzazione il riconoscimento di parole brevi o caratteri speciali per i quali il livello di accuratezza non è elevato. Unitamente alle parole, il sistema procede al riconoscimento e alla classificazione dei simboli attraverso un modello di Symbol Prediction Fully Convolutional Network addestrato, nel caso specifico, su una base di conoscenza composta da 110 simboli e caratteri speciali.

Al giorno d'oggi sono stati implementati un congruo numero di strumenti e software, taluni disponibili open source e altri a pagamento, in grado di elaborare il riconoscimento della scrittura con un certo grado di accuratezza. Transkribus è certamente tra le applicazioni, basate sull'intelligenza artificiale per l'ICR, più adoperate al mondo grazie alla qualità dei risultati e alla semplicità di utilizzo. Una delle caratteristiche principali di Transkribus, è la possibilità di addestrare ulteriormente l'applicativo e, di conseguenza, incrementare il livello di accuratezza sulla scrittura presa in oggetto. La fase di addestramento implica la creazione, da parte degli utenti, della base di conoscenza, ovvero le trascrizioni su un numero adeguato di pagine necessarie per l'addestramento del modello di riconoscimento della scrittura. Il processo è mediato da un'interfaccia grafica di semplice utilizzo che non richiede alcuna abilità di programmazione, attraverso la quale l'utente è in grado di fornire la

trascrizione corrispondente alla parte di testo selezionata. Inoltre, la facilità attraverso cui l'addestramento può essere effettuato favorisce la collaborazione tra utenti che, in tal modo, contribuiscono ad alimentare la base di conoscenza dell'applicativo e a renderlo accurato su un numero maggiore di scritture. Sulla base di tali caratteristiche, Transkribus si configura come una soluzione indirizzata, particolarmente, verso la comunità scientifica e accademica, per loro natura maggiormente propense verso la collaborazione. Esistono anche soluzioni ideate per scopi commerciali unendo strumenti di OCR e ICR di elevata precisione, es. FlexiCapture Engine o A2IA Mitek , o soluzioni interamente gratuite come le librerie Python, ad es. ICR e *Intelligent-Character-Recognition*.

Nel settore culturale la scrittura costituisce una fonte informativa importante, nonché una forma di espressione di eccezionale rilevanza storica, culturale e artistica. Come tale, l'applicazione di tecniche di ICR consente di capitalizzare la conoscenza, espressa in forma testuale e grafica come ad esempio appunti o composizioni artistico-letterarie, potenzialmente utili a delineare nuove e diverse chiavi interpretative.

Audio e trattamento di audiovideo

Il riconoscimento automatico del parlato – *Automatic Speech Recognition (ASR)* – è un campo connesso alla computer science e alla linguistica computazionale con l'obiettivo di sviluppare tecnologie per l'estrazione e la trascrizione del parlato¹⁸⁶. Oggetto di studio è il segnale sonoro che viene processato per individuare una fonte orale da “isolare” rispetto al contesto per fornirne una rappresentazione testuale. Si tratta di un processo molto complesso la cui buona riuscita dipende da diversi fattori quali, ad esempio, il tono o il timbro di voce di ciascun individuo, lo stato emotivo, la chiarezza espositiva, lo stile etc.

Le prime tecniche di *speech recognition* nascono a partire tra la fine del '50 e l'inizio degli anni '60 dapprima con l'implementazione di *Audrey*¹⁸⁷ del Bell Labs, un sistema in grado di riconoscere l'enunciazione di numeri, e successivamente con il modello *source-filter* di Gunnar Fant¹⁸⁸ e il 16-

¹⁸⁶ Trabelsi, Asma, Sébastien Warichet, Yassine Aajaoun, and Séverine Soussilane. “Evaluation of the efficiency of state-of-the-art Speech Recognition engines.” *26th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems*. 2022. p. 1. Doi:/10.1016/j.procs.2022.09.534.

¹⁸⁷ Soni, Vishal Dineshkumar. “Speech Recognition: Transcription and transformation of human speech.” *International Journal on Integrated Education* 2, n. VI. 2019. p. 2. <https://doi.org/10.31149/ijie.v2i6.497>.

¹⁸⁸ Soni, Vishal Dineshkumar. “Speech Recognition: Transcription and transformation of human speech.” *International Journal on Integrated Education* 2, n. VI. 2019. p. 2. <https://doi.org/10.31149/ijie.v2i6.497>.

word shoebox della IBM¹⁸⁹. Attualmente è possibile distinguere tre principali approcci di *speech recognition*¹⁹⁰ (Figura 19):

1. *Acoustic Phonetic Approach*. Presuppone una conoscenza di base degli aspetti fonetici e si basa sull'individuazione delle caratteristiche specifiche del parlato tramite le onde sonore prodotte dagli organi vocali umani. E' un approccio che viene utilizzato per implementare le regole fonetiche nei sistemi di riconoscimento vocale.
2. *Pattern Recognition Approach*. Questo approccio non richiede una conoscenza a priori del parlato e si costituisce di due fasi: (i) addestramento di modelli vocali attraverso determinati comuni set di parametri spettrali e (ii) rilevamento di modelli mediante l'analisi del modello. Questo approccio risulta molto performante richiedendo una bassa potenza di calcolo. Alcune tecniche di esempio sono il *Support Vector Machine (SVC)* e l'*Hidden Markov Model (HMM)*.
3. *Artificial Intelligence Approach*. Costituisce un approccio ibrido tra l'*acoustic phonetic* e il *pattern recognition* e impiega tecniche di *Deep Learning* per l'analisi degli spettrogrammi da cui estrarre le caratteristiche del parlato.

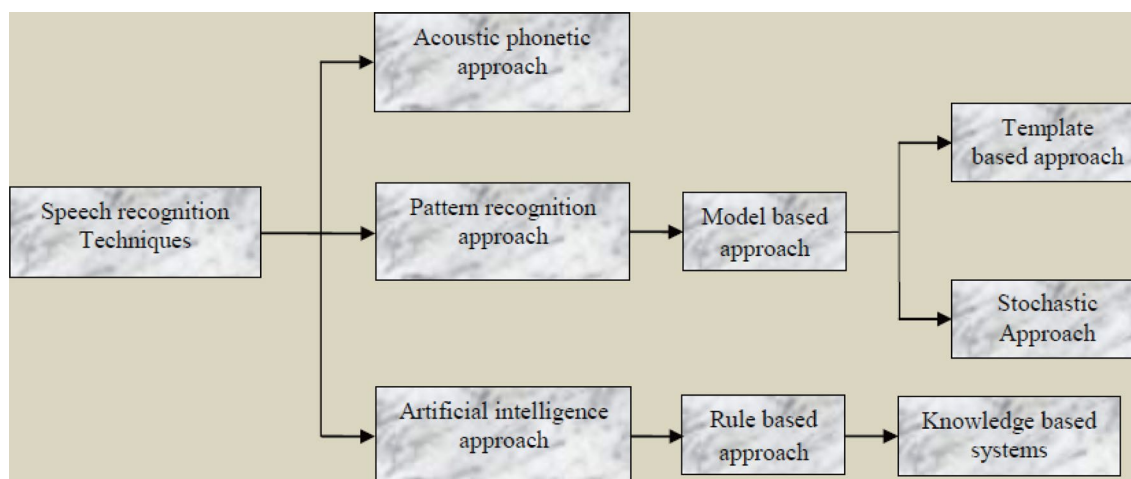


Figura 19: Classificazione dei sistemi di riconoscimento del parlato.¹⁹¹

In base alla tipologia di operazione effettuata, i sistemi di riconoscimento del parlato possono essere classificati in base a:

- Enunciato

¹⁸⁹ Soni, Vishal Dineshkumar. "Speech Recognition: Transcription and transformation of human speech." *International Journal on Integrated Education* 2, n. VI. 2019. p. 2. <https://doi.org/10.31149/ijie.v2i6.497>.

¹⁹⁰ Vaishali, Kherdekar A., and Sachin A. Naik. "Speech Recognition System Approaches, Techniques And Tools For Mathematical Expressions: A Review." *International Journal of Scientific & Technology Research* 8, n. 8. 2019. p. 1.

¹⁹¹ Haridasa, Arul Valiyavalappil, Ramalatha Marimuthub, and Vaazi Gangadharan Sivakumarc. "A critical review and analysis on techniques of speech recognition: The road ahead." *International Journal of Knowledge-based and Intelligent Engineering Systems*, 22, n. 1. 2018, p. 4. DOI: 10.3233/KES-180374.

- Parlante
- Dimensione del vocabolario.

All'interno della prima categoria è possibile distinguere ulteriormente quattro tipologie diverse a seconda che si tratti di (i) Parole isolate, per riconoscere parole pronunciate singolarmente come nei casi di vocabolari; (ii) Parole connesse, in cui è richiesta una pausa breve tra l'enunciazione di parole; (iii) Discorso continuo, viene permesso all'utente di parlare con maggiore naturalezza componendo frasi intere; (iv) Discorso spontaneo, l'utente pronuncia discorsi in piena autonomia e questo può includere l'uso di parole prive di significati, interruzioni o balbuzie.

Relativamente al parlante, questa categoria distingue diversi gradi di dipendenza del singolo utente:

- Piena dipendenza: il modello è costruito per riconoscere il parlato da un utente specifico. È più semplice da realizzare ma è accurato per il singolo caso.
- Indipendente dal parlante: il modello funziona su utenti generici. Più complesso da implementare con un'accuratezza generica.
- Adattamento al parlante: modello costruito per adattarsi ai nuovi utenti. L'accuratezza aumenta gradualmente.

Infine, l'ultima classe è relativa al vocabolario, classificato in base alla grandezza, distinguendo: (i) piccolo, 10 parole; (ii) medio, 11-100 parole; (iii) grande, circa 1000; (iv) molto grande, circa 10.000; (v) al di là del dizionario, sono incluse parole non presenti nel dizionario.

Il processo di ASR ha inizio con la creazione di sequenze di *features vector* acustici dal segnale sonoro. Questa attività comporta, dunque, la preliminare conversione dall'analogico al digitale del segnale sonoro da cui procedere per estrarne, per ogni segmento di parlato o *frame*, le caratteristiche spettrali. Si tratta, in questo caso, di un'operazione di primaria importanza da cui dipende il grado di accuratezza del risultato finale. Il processo di estrazione delle caratteristiche implica l'impiego di tecniche per ricavare specifici parametri dall'enunciazione del parlato nei vettori di informazione. Alcune di queste tecniche sono¹⁹²:

- *Mel Frequency Cepstral Coefficient* (MFCC). Si tratta della tecnica di maggior applicazione adoperata per il riconoscimento multiparlante, multilingua e con vocabolari di grandi dimensioni.
- *Linear Predictive Coding* (LPC). È un metodo statico per l'estrazione delle caratteristiche in forma di intervallo di *pitch*¹⁹³, energia del parlato per frame etc. Funziona soprattutto in contesti di assenza di rumore in sottofondo.

¹⁹² Vaishali, Kherdekar A., and Sachin A. Naik. *Cit.* p. 4.

¹⁹³ "la sensazione uditiva che consente di ordinare un suono su una scala che si estenda dal basso verso l'alto. Il pitch dipende principalmente dal contenuto di frequenze dello stimolo sonoro, ma anche dalla pressione e dalla forma dello stimolo stesso." Treccani. "Pitch." www.treccani.it/enciclopedia/pitch/.

- *Relative Spectral Transform* (RASTA Filtering). È la tecnica impiegata ~~in particolare~~ nei casi in cui vi sia rumore in sottofondo. Impiega un filtro che accetta le frequenze all'interno di un determinato range prestabilito, rigettando quelle che si collocano al di fuori.

Alle tecniche di estrazione delle *features*, si aggiunge il modello per espletare il task di riconoscimento e trascrizione del parlato. Genericamente, è possibile distinguere due tipologie di architetture: una basata su un sistema multi modulo e un'altra su un modello *end-to-end*. Nel primo caso, come da Figura 20, le operazioni sono distinte le une dalle altre implementando singolarmente i rispettivi moduli. Dal task di *language models* vengono prodotti i modelli linguistici di riferimento con cui costruire il set di parole a cui applicare i modelli fonetici (*pronunciation modules*). A partire da quest'ultimi saranno generate sequenze di token di pronuncia come input per i modelli acustici (*acoustic models*), dai quali saranno prodotti i corrispettivi suoni. È il caso dei modelli basati su *Hidden Markov Model* (HMM) che hanno rappresentato per anni il modello di riferimento, in particolare nelle elaborazioni su vocabolari di grandi dimensioni.

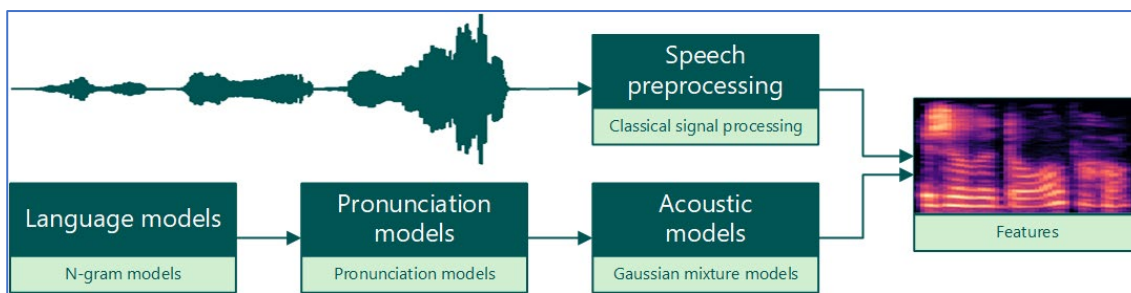


Figura 20: sistema di riconoscimento automatico del parlato multi modulo.¹⁹⁴

Nel caso dei sistemi *end-to-end*, queste elaborazioni sono implementate come parte integrante di un unico modello e, quindi, non più come singoli moduli. Genericamente, la maggior parte di questi sistemi si compone di *encoder* e *decoder* (Figura 21).

¹⁹⁴ Iosifova, Olena, Ievgen Iosifov, Volodymyr Sokolov, Oleh Romanovskyi, and Igor Sukaylo. "Analysis of Automatic Speech Recognition Methods". *CEUR Workshop Proceedings*. 2021.

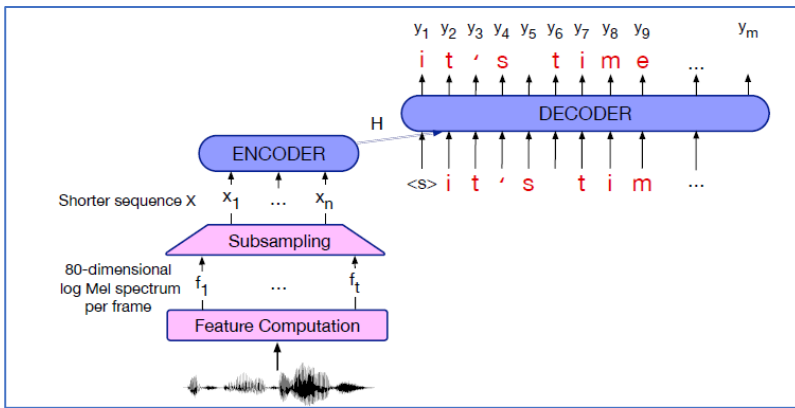


Figura 21: Schema dell'architettura ASR basata su encoder-decoder.¹⁹⁵

Il punto di partenza è la sequenza di *feature vectors* che viene processata dalla componente di *encoder* per restituire come risultato, tramite il *decoder*, lettere e numeri secondo il modello linguistico associatovi. Questa architettura richiede un preliminare livello di compressione (*Subsampling*) in cui ridurre la sequenza delle caratteristiche acustiche da dare in input all'*encoder*. I principali esempi di questa classe di sistemi sono i modelli *Connection Temporal Classification (CTC)* e *Recurrent Neural Network-Transducer*. CTC mappa le sequenze di vettori in input nella sequenza di etichette in output mediante due sottoprocessi: *path probability calculation* e *path aggregation*. Nella prima parte, si procede a calcolare la probabilità che un elemento processato della sequenza di *feature vector* appartenga ad una precisa etichetta del vocabolario di riferimento. Il risultato di questa fase è una sequenza di trascrizione lunga quanto la sequenza acustica in input. Dal momento che, generalmente, la lunghezza della trascrizione risulta essere ben più breve di quella acustica¹⁹⁶, nella fase di *path aggregation* viene applicato un *mapping* per stabilire la trascrizione finale del parlato. Uno dei maggiori limiti di questo modello si rivela nei casi in cui la lunghezza della sequenza di trascrizione risulta essere maggiore rispetto a quella del parlato, per i quali il modello risulta non essere molto efficace. Il modello RNN-Transducer si differenzia da CTC nei metodi di elaborazione che permettono di superare alcuni limiti riscontrabili in quest'ultimo. RNN-Transducer consiste di tre sottoreti addestrate congiuntamente:

1. *Transcription network*. Rappresenta il codificatore con funzioni di modello acustico.
2. *Prediction network*. È parte dell'elemento di decodifica con funzioni di modello linguistico.
3. *Joint network*. Elabora l'allineamento tra le sequenze di input e di output.

¹⁹⁵ Jurafsky, Daniel, and James H. Martin. *Speech and Language Processing*

An Introduction to Natural Language Processing. Computational Linguistics, and Speech Recognition. 2023, p. 337.

¹⁹⁶ Wang, Deng, Xiaodong Wang, and Shaohe Lv. "An Overview of End-to-End Automatic Speech Recognition." *Symmetry*, 11. 2019, p. 7. doi:10.3390/sym11081018.

Prediction e Joint sono correlate tra loro in modo che la previsione successiva sia dipendente da quella precedente. L'interazione dei moduli, addestrati congiuntamente, rende superfluo ricorrere a un modello linguistico esterno.

Rispetto al modello CTC, RNN-Transducer permette la generazione in output di sequenze di lunghezza arbitraria indipendentemente che esse siano maggiori o inferiori di quelle in input. Di contro, ogni componente di questo modello richiede un pre-addestramento per ottenere buoni risultati.

Part of Speech Tagging

Il *Part of Speech Tagging* (POS) – Identificazione delle Parti del Discorso – è il task di classificazione delle parole in una frase in base alla funzione sintattica che esse assumono. Individuare il ruolo di un termine all'interno di una frase contribuisce a determinare il significato assunto dalle parole adiacenti e, così facendo, risolvere eventuali ambiguità. Si basa sui costrutti sintattici e, come tale, è fortemente condizionato dalla lingua a cui si applica. Sono stati sviluppati diversi dataset per l'addestramento dei modelli di POS Tagging in base alla lingua impiegata, o anche multilingue. Nel caso dell'italiano, ad esempio, è stato sviluppato il corpus TINT (*The Italian NLP Tool*)¹⁹⁷ elaborato traducendo e adattando il dataset multilingua¹⁹⁸ Stanford CoreNLP¹⁹⁹.

Dataset Audio

Il riconoscimento automatico del parlato può essere applicato in molteplici contesti e per scopi ben precisi come, ad esempio, lo sviluppo di assistenti digitali, traduttori automatici, chatbot e via di seguito. In base agli scopi prefissati, alcune tipologie di dataset risulteranno più congrue per l'addestramento dei sistemi.

Allo stato attuale sono stati creati innumerevoli dataset che raccolgono centinaia di ore di discorsi per addestrare i vari modelli al task di riconoscimento automatico del parlato²⁰⁰. È possibile procedere ad una loro classificazione in base a:

- lingua: se prevedono una lingua o dialetto specifico oppure essere multilingua;
- argomento: discorsi pronunciati in un dato contesto o su un tema preciso;
- rapporto tra i parlanti: interazioni in contesti differenti come situazioni ufficiali, persone conoscenti o estranei etc.

¹⁹⁷Apro시오, Alessio Palmero, and Giovanni Moretti. "Tint 2.0: an All-inclusive Suite for NLP in Italian." *Proceedings of the Fifth Italian Conference on Computational Linguistics CLIC-it*, 10. 2018.

¹⁹⁸ Allo stato attuale supporta otto lingue: Arabo, Cinese, Inglese, Francese, Tedesco, Ungherese, Italiano e Spagnolo.

¹⁹⁹ Manning, Christopher D., Mihai Surdeanu, John Bauer, Jenny Finkel, Steven J. Bethard, and David McClosky. "The Stanford CoreNLP Natural Language Processing Toolkit." *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*. 2014.

²⁰⁰ Per maggiori dettagli circa il numero e la varietà di dataset disponibili, si suggerisce di visitare il repository: <https://paperswithcode.com/datasets?task=speech-recognition>.

Di seguito sono presentati alcuni tra i dataset maggiormente impiegati:

- LibriSpeech.²⁰¹ Contiene oltre 1000 ore di audiolibri estratti dal progetto LibriVox, suddiviso, per la parte di training, nelle classi (i) 100h, (ii) 360h e (iii) 500h. La parte di dati adibiti per il test è invece distinta in “clean”, più semplice da processare, e “other”, in base al grado di difficoltà con cui addestrare i sistemi di ASR.
- Switchboard.²⁰² Corpus composto da oltre 200 ore di conversazioni telefoniche collezionate nei primi anni '90 utile, in particolare, per l'addestramento di sistemi di chatbot telefonici.
- VoxPopuli.²⁰³ Si tratta di un corpus particolarmente ricco sia sotto il profilo delle ore – circa 100.000 – e sia per il numero di lingue – 23 – in cui sono presenti discorsi non etichettati. Sono inoltre presenti 1800 ore di trascrizioni in 16 differenti lingue.
- AISHELL-1.²⁰⁴ È un corpus definito specificatamente per i sistemi ASR per il mandarino. Contiene oltre 170 ore di letture afferenti vari domini e con oratori provenienti da diverse zone della Cina.

Strumenti di estrazione automatica delle informazioni applicati alle immagini

Il progresso compiuto nel campo del Deep Learning ha portato allo sviluppo di una mole di algoritmi in grado di processare le immagini sotto diversi punti di vista come la classificazione (*Image classification*), l'individuazione e segmentazione di oggetti (*Object detection* e *Object segmentation*) o la generazione automatica di altre immagini (*Image generation*). Per le finalità del caso di studio riportato in questo elaborato, si prenderanno in esame le tecniche e i modelli di DL per le attività di classificazione e riconoscimento degli oggetti applicati per estrarre contenuti all'interno delle immagini.

Trattandosi di attività con un elevato grado di complessità che richiedono l'implementazione di modelli articolati per ottenere risultati soddisfacenti, sono stati presi a riferimento modelli precostituiti di reti neurali convolutive che possono essere applicati con il loro addestramento di base, oppure essere nuovamente addestrati su altri dataset. Per il caso di studio in oggetto sono state prese in esame le principali architetture di CNN rinvenibili in letteratura scientifica i cui modelli sono disponibili tramite i framework Keras e PyTorch. In aggiunta a questo criterio i modelli sono stati

²⁰¹ Han, Kyu J., Ramon Prieto, Kaixing Wu, and Tao Ma. “State-of-the-Art Speech Recognition Using Multi-Stream Self-Attention With Dilated 1D Convolutions.” 2019. arXiv:1910.00716.

²⁰² Godfrey, John J., and Edward Holliman. “Switchboard-1 Release 2 LDC97S62.” 1993. Doi:10.35111/sw3h-rw02.

²⁰³ Wang, Changhan, Morgane Riviere, Ann Lee, Anne Wu, Chaitanya Talnikar, Daniel Haziza, Mary Williamson, Juan Pino, Emmanuel Dupoux. “VoxPopuli: A Large-Scale Multilingual Speech Corpus for Representation Learning, Semi-Supervised Learning and Interpretation.” 2021. arXiv:2101.00390v2.

²⁰⁴ Bui, Hu, Jiayu Du, Xingyu Na, Bengu Wu, and Hao Zheng. “AISHELL-1: An open-source mandarin speech corpus and a speech recognition baseline.” 2017. arXiv:1709.05522v1.

selezionati prendendo in considerazione la tabella comparativa disponibile in Keras²⁰⁵, dalla quale sono selezionati:

- Dimensioni del modello: misura in Megabyte;
- Top-1 e Top-5 Accuracy: misura di accuratezza per l'identificazione del singolo oggetto principale (Top-1) e dei primi cinque elementi riscontrabili nell'immagine (Top-5);
- Numero di parametri: il numero di parametri disposti in input nel modello che contribuiscono a fornire maggiore capacità di calcolo alla rete rendendola, di contro, più “pesante” sotto il profilo computazionale;
- Profondità: indica il numero di strati presenti nel modello;

Come specificato in Keras, questi parametri sono dettati dai risultati ottenuti sul *validation dataset* di ImageNet impiegando una CPU AMD con processore EPYC, RAM da 1,7 T, GPU Tesla A100 e una dimensione del Batch impostata a 32 esempi.

Model	Size (MB)	Top-1 Accuracy	Top-5 Accuracy	Parameters	Depth
Xception	88	79.0%	94.5%	22.9M	81
VGG16	528	71.3%	90.1%	138.4M	16
VGG19	549	71.3%	90.0%	143.7M	19
ResNet50	98	74.9%	92.1%	25.6M	107
ResNet50V2	98	76.0%	93.0%	25.6M	103
ResNet101	171	76.4%	92.8%	44.7M	209
ResNet101V2	171	77.2%	93.8%	44.7M	205
ResNet152	232	76.6%	93.1%	60.4M	311
ResNet152V2	232	78.0%	94.2%	60.4M	307
InceptionV3	92	77.9%	93.7%	23.9M	189
InceptionResNetV2	215	80.3%	95.3%	55.9M	449
MobileNet	16	70.4%	89.5%	4.3M	55
MobileNetV2	14	71.3%	90.1%	3.5M	105
DenseNet121	33	75.0%	92.3%	8.1M	242
DenseNet169	57	76.2%	93.2%	14.3M	338
DenseNet201	80	77.3%	93.6%	20.2M	402
NASNetMobile	23	74.4%	91.9%	5.3M	389
NASNetLarge	343	82.5%	96.0%	88.9M	533
EfficientNetB0	29	77.1%	93.3%	5.3M	132
EfficientNetB1	31	79.1%	94.4%	7.9M	186
EfficientNetB2	36	80.1%	94.9%	9.2M	186
EfficientNetB3	48	81.6%	95.7%	12.3M	210
EfficientNetB4	75	82.9%	96.4%	19.5M	258
EfficientNetB5	118	83.6%	96.7%	30.6M	312
EfficientNetB6	166	84.0%	96.8%	43.3M	360
EfficientNetB7	256	84.3%	97.0%	66.7M	438

²⁰⁵ Keras, “API Applications,” ultimo accesso 23 Gennaio 2024, <https://keras.io/api/applications/>.

Model	Size (MB)	Top-1 Accuracy	Top-5 Accuracy	Parameters	Depth
EfficientNetV2B0	29	78.7%	94.3%	7.2M	-
EfficientNetV2B1	34	79.8%	95.0%	8.2M	-
EfficientNetV2B2	42	80.5%	95.1%	10.2M	-
EfficientNetV2B3	59	82.0%	95.8%	14.5M	-
EfficientNetV2S	88	83.9%	96.7%	21.6M	-
EfficientNetV2M	220	85.3%	97.4%	54.4M	-
EfficientNetV2L	479	85.7%	97.5%	119.0M	-
ConvNeXtTiny	109.42	81.3%	-	28.6M	-
ConvNeXtSmall	192.29	82.3%	-	50.2M	-
ConvNeXtBase	338.58	85.3%	-	88.5M	-
ConvNeXtLarge	755.07	86.3%	-	197.7M	-
ConvNeXtXLarge	1310	86.7%	-	350.1M	-

Tabella 6: Tabella di comparazione tra i modelli di CNNs disponibili in Keras²⁰⁶.

All'interno di Bianco et al. (2018)²⁰⁷, è stato condotto uno studio di rilievo sulla comparazione e misurazione dell'efficienza tra le principali architetture di reti neurali convolutive, quali ad esempio: AlexNet, la famiglia delle VGG, BN-Inception, GoogLeNet, SqueezeNet v.1.0 e v.1.1, ResNet (nelle versioni -18, -34, -50, -101 e -152), DenseNet (nelle versioni 121, -169 e -201)²⁰⁸. La sperimentazione è stata compiuta impiegando il *validation set* di ImageNet-1k come metro di paragone. L'accuratezza, espressa nelle misure di Top-1 e Top-5, è stata valutata tenendo conto della complessità computazionale, calcolata in *Floating-point operations* (FLOPs), e dell'articolazione del modello in numero di strati e di parametri. I grafici nella Figura 22²⁰⁹ mostrano i risultati ottenuti da questa sperimentazione.

²⁰⁶ Keras, "API Applications," ultimo accesso 23 Gennaio 2024, <https://keras.io/api/applications/>.

²⁰⁷ Simone Bianco, Remi Cadene, Luigi Celona, and Paolo Napoletano, "Benchmark Analysis of Representative Deep Neural Network Architectures," (2018). arXiv:1810.00736.

²⁰⁸ Per una lista completa si rimanda a Simone Bianco, Remi Cadene, Luigi Celona, and Paolo Napoletano. *cit.* p.2.

²⁰⁹ Simone Bianco, Remi Cadene, Luigi Celona, and Paolo Napoletano. *cit.* p.3.

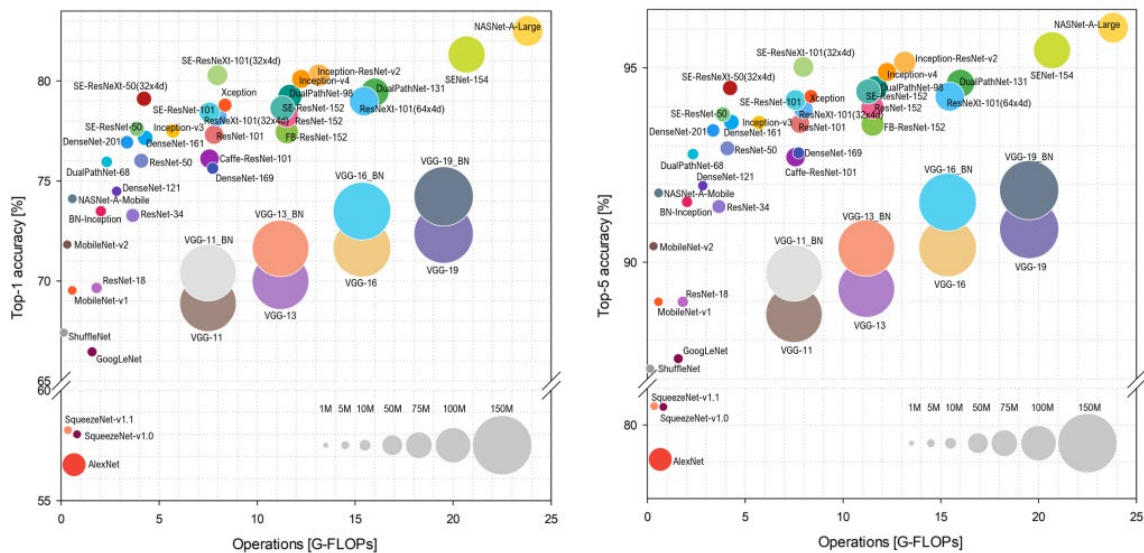


Figura 22: Risultati Top-1 e Top-5 vs computazionale complessità.

All'interno dei grafici le dimensioni delle sfere rappresentano il grado complessità di ciascun modello, sull'asse x sono riportati i valori FLOPs mentre sull'asse y il grado di accuratezza raggiunto, rispettivamente Top-1 per il grafico a sinistra e Top-5 per quello a destra. Quanto emerge da questo lavoro testimonia la non correlazione tra grado di complessità dei modelli e accuratezza raggiunta. Difatti, taluni modelli poco complessi come Xception, SE-ResNetXt-101 o SE-ResNet-101, hanno ottenuto risultati elevatissimi con costi computazionali contenuti. Nel caso del VGG-13, la cui complessità è testimoniata dalle dimensioni della sfera, l'accuratezza raggiunta è a livello di quella ottenuta da ResNet-18 le cui dimensioni sono notevolmente inferiori.

In merito alla comparazione tra i modelli sono stati compiuti vari lavori come nel caso di Darapaneni et al (2020)²¹⁰, in cui l'analisi è stata compiuta specificatamente per l'attività di *Image classification*, testando i modelli sul dataset di ImageNet. Sono stati avanzati lavori comparativi anche in merito a tematiche specifiche come nel caso in Du et al. (2023)²¹¹ dove i modelli AlexNet, VGG-16, GoogLeNet e ResNet sono stati sperimentati sul dataset sottomarino Side-scan Sonar (SSS) contenente immagini di aeroplani e navi, o in Çinar et al. (2022)²¹² dove i modelli DenseNet169, AlexNet, InceptionV3 e ResNet101 sono stati applicati per l'individuazione di tumori al cervello da immagini a risonanza magnetica.

²¹⁰ Narayana Darapaneni, Krishnamurthy B, and Anwesh Reddy Paduri, "Convolution Neural Networks: A Comparative Study for Image Classification," *15th (IEEE) International Conference on Industrial and Information Systems (ICIIS)*, (2020). DOI: 10.1109/ICIIS51140.2020.9342667.

²¹¹ Xing Du, Yongfu Sun, Yupeng Song, Huifeng Sun, and Lei Yang, "A Comparative Study of Different CNN Models and Transfer Learning Effect for Underwater Object Classification in Side-Scan Sonar Images," *Remote Sensing* 15, n. 3: 593. DOI:10.3390/rs15030593.

²¹² Necip Çinar, Kaya Buket, and Kaya Mehemet. "Comparison of deep learning models for brain tumor classification using MRI images," *2022 International Conference on Decision Aid Sciences and Applications (DASA)*, (2022). DOI: 10.1109/DASA54658.2022.9765250.

Nella fase di scelta degli strumenti, lo studio di questi lavori è stato indispensabile per avere evidenza delle caratteristiche e delle potenzialità dei diversi modelli per procedere ad una scelta ponderata per il caso di studio. Sulla base delle informazioni derivanti dai lavori di comparazione, è stato necessario commisurare la scelta di quale o più modelli applicare per il task di *Object detection e classification* tenendo in considerazione la strumentazione e la potenza di calcolo disponibile per l'attività di ricerca. Infine, si è proceduto a verificare la disponibilità di tali modelli all'interno dei framework Keras e Pytorch. La tabella seguente illustra una prima selezione dei modelli indicando per ognuno di essi il grado di accuratezza testimoniato.

MODELLO	Top-1 Accuracy (%)	Top-5 Accuracy (%)	Keras	PyTorch	Parameters
Xception	75-80	90-95	Si	No	22.9M
VGG-16	70-75	90-93	Si	Si	138.4M
ResNet101	75-80	93-95	Si	Si	44.7M
Inception-v3	75-80	93-95	Si	Si	23.9M
NasNet-A-Large	>80	>95	Si	No	88.9M
ResNet-152	77-80	93-95	Si	Si	60.4M
DenseNet-169	75-77	90-93	Si	Si	14.3M
VGG-19	70-75	90-93	Si	Si	143.7M

Tabella 7: modelli selezionati in fase iniziale.

Nella fase successiva, a partire da questa prima selezione, sono stati scelti i modelli con un rapporto accuratezza - numero di parametri adeguato alla strumentazione disponibile. Dai modelli elencati in Tabella 7, sono stati selezionati i modelli Xception, ResNet152 e NasNet-A-Large.

Xception²¹³ è stata sviluppata all'interno di Google come una interpretazione dell'architettura Inception basata su una struttura composta di strati di convoluzione separabili in profondità. Questa architettura prevede 36 strati di convoluzione, sui quali si basa l'attività di *features extraction* della rete, articolati in 14 moduli tra i quali sono stabilite delle connessioni lineari, fatta eccezione per lo strato iniziale e finale. Per la somiglianza nel numero di parametri e architettura, in fase sperimentale questa architettura è stata confrontata con Inception V3 sulle attività di classificazione con singolo oggetto e con oggetti multipli scegliendo come dataset di riferimento ImageNet per la prima attività e JFT per la seconda. In entrambe sperimentazioni Xception si è rivelata migliore rispetto a Inception v3 ottenendo nel primo caso il 79% per Top-1 e 94,5% per Top-5 accuracy contro il rispettivo 78,2%

²¹³ François Chollet, "Xception: Deep Learning with Depthwise Separable Convolutions," (2016). <https://arxiv.org/abs/1610.02357>.

e 94,1% di Inception v3. Sul dataset di JFT il margine sale al 4,3% con la metrica Mean Average Precision for top 100 predictions (MAP@100). Di seguito è riportata l'architettura di Xception²¹⁴.

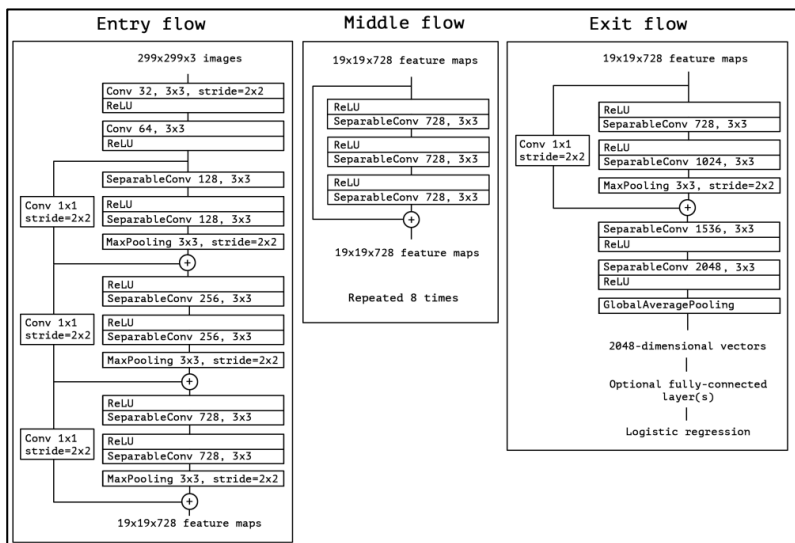


Figura 23: Struttura dell'architettura di Xception²¹⁵.

ResNet152²¹⁶ è una versione di CNN composta da 152 strati basata sulla tecnica di “skip connections”, da cui Residual Network, che consente di mitigare al problema della perdita di gradienti, nelle reti con molti strati, durante la fase di retropropagazione permettendo un flusso costante dei gradienti tra gli strati della rete. Questa tecnica connette uno strato con gli altri saltandone alcuni in posizione intermedia. Gli strati saltati formano tra loro un blocco residuo (*residual block*) e, nel caso specifico di ResNet101, tre blocchi definiti *Bottleneck* all'interno dei quali vengono svolte operazioni di convoluzione e normalizzazione. Ognuno di questi blocchi è attivato individualmente riducendo in tal modo il numero di parametri da processare e ottimizzare. Allo stato attuale sono state implementate diverse estensioni dell'architettura ResNet come riportato nella figura²¹⁷ sottostante in cui è illustrata l'architettura per le versioni a 18, 34, 50, 101 e 152 strati.

²¹⁴ François Chollet, “Xception: Deep Learning with Depthwise Separable Convolutions,” cit. p.5.

²¹⁵ François Chollet, “Xception: Deep Learning with Depthwise Separable Convolutions,” cit. p.5.

²¹⁶ Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. “Deep Residual Learning for Image Recognition,” (2015). arXiv:1512.03385v1.

²¹⁷ Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. “Deep Residual Learning for Image Recognition,” cit. p.5.

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112	7×7, 64, stride 2				
		3×3 max pool, stride 2				
conv2_x	56×56	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
conv3_x	28×28	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 8$
conv4_x	14×14	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 23$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 36$
conv5_x	7×7	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
	1×1	average pool, 1000-d fc, softmax				
FLOPs		1.8×10 ⁹	3.6×10 ⁹	3.8×10 ⁹	7.6×10 ⁹	11.3×10 ⁹

Figura 24: Architettura di ResNet²¹⁸.

NasNetLarge²¹⁹ è una rete caratterizzata dal processo di Neural Architecture Search (NAS) definito per l'ottimizzazione delle operazioni di convoluzione su dataset di grandi dimensioni. Questo concetto permette al modello di apprendere a partire dal dataset di interesse senza che la dimensione di questi influisca sulle performance di apprendimento. Peculiarità di questa rete sono le Normal Cells e Reduction Cells, celle convolutive finalizzate, le prime, a restituire una griglia delle caratteristiche invariata e, le seconde, a ridurre le dimensioni della griglia in input. Non essendo predeterminata la struttura delle celle convolutive, questa viene predetta a partire dagli strati iniziali nascosti applicando un controller di una rete neurale ricorsiva (RNN) e, specificatamente, un'architettura LSTM. Questa operazione è finalizzata a individuare la migliore struttura per la cella partendo da un set di operazioni (Figura 25). Da questo processo²²⁰ sono state generate tre tipologie diverse di modelli: NASNet-A, NASNet-B e NASNet-C. Tra questi, il modello che ha ottenuto i migliori risultati, e per tale motivo oggetto di studio in questo lavoro di ricerca, è NESNet-A di cui sono riportati nella Figura 18 i risultati ottenuti sull'attività di *object detection* sul dataset COCO.

²¹⁸ Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep Residual Learning for Image Recognition," *cit.* p.5.

²¹⁹ Barret Zoph, Vijay Vasudevan, Honathon Shlens, and Quoc V.Le. "Learning Transferable Architectures for Scalable Image Recognition," (2018). arXiv:1707.07012v4.

²²⁰ Per i dettagli sulla durata del processo e sugli strumenti si veda Barret Zoph, Vijay Vasudevan, Honathon Shlens, and Quoc V.Le. *cit.* pp.4-7.

- | | |
|--------------------------------|--------------------------------|
| ● identity | ● 1x3 then 3x1 convolution |
| ● 1x7 then 7x1 convolution | ● 3x3 dilated convolution |
| ● 3x3 average pooling | ● 3x3 max pooling |
| ● 5x5 max pooling | ● 7x7 max pooling |
| ● 1x1 convolution | ● 3x3 convolution |
| ● 3x3 depthwise-separable conv | ● 5x5 depthwise-separable conv |
| ● 7x7 depthwise-separable conv | |

Figura 25: Set di operazioni da selezionare con RNN²²¹.

Model	resolution	mAP (mini-val)	mAP (test-dev)
MobileNet-224 [24]	600 × 600	19.8%	-
ShuffleNet (2x) [70]	600 × 600	24.5% [†]	-
NASNet-A (4 @ 1056)	600 × 600	29.6%	-
ResNet-101-FPN [36]	800 (short side)	-	36.2%
Inception-ResNet-v2 (G-RMI) [28]	600 × 600	35.7%	35.6%
Inception-ResNet-v2 (TDM) [52]	600 × 1000	37.3%	36.8%
NASNet-A (6 @ 4032)	800 × 800	41.3%	40.7%
NASNet-A (6 @ 4032)	1200 × 1200	43.2%	43.1%
ResNet-101-FPN (RetinaNet) [37]	800 (short side)	-	39.1%

Figura 26: Performance relative all'object detection sul dataset COCO²²².

Strumenti di estrazione automatica delle informazioni applicati agli audio-video

Nel caso dei documenti audiovisivi è necessario dapprima separare le due componenti principali: la fonte sonora e le immagini dinamiche. Quest'ultime, similmente alle immagini statiche, possono essere processate tramite le architetture di CNN per le operazioni di classificazione e riconoscimento degli oggetti. Ciò che maggiormente caratterizza i video, tuttavia, è la possibilità di addestrare dei modelli per il riconoscimento delle attività manifestate all'interno. Per la componente sonora, invece, è possibile applicare le reti neurali per l'estrazione e la trascrizione del parlato cosicché da rendere possibile l'analisi tramite NLP. Per le operazioni di *video classification* e *activity recognition* è possibile impiegare modelli precostituiti da addestrare sul dataset UCF101²²³, composto da 13.320 video suddivisi in 101 classi di attività. Come modello di riferimento si prenda InceptionV3, suggerito all'interno della guida e tutoria di Keras²²⁴ per il task di *video classification* e *activity recognition*.

²²¹ Per i dettagli sulla durata del processo e sugli strumenti Cfr. Barret Zoph, Vijay Vasudevan, Honathon Shlens, and Quoc V.Le. *cit.* p.3.

²²² Barret Zoph, Vijay Vasudevan, Honathon Shlens, and Quoc V.Le. *cit.* p.3.

²²³ University of Central Florida, "UFC Dataset," ultimo accesso 23 Gennaio 2024, <https://www.crcv.ucf.edu/data/UCF101.php>.

²²⁴ Keras, "Video Classification," ultimo accesso 23 Gennaio 2024, https://keras.io/examples/vision/video_classification/.

InceptionV3²²⁵ è un modello di rete neurale convoluzionale composto di 42 strati noto per l'efficace classificazione delle immagini. Rispetto alla versione precedente, InceptionV1, sono stati introdotti alcuni miglioramenti in grado di rendere il nuovo modello più efficiente con minori costi computazionali adottando un'architettura che utilizza strati paralleli per prevenire l'*overfitting*. Le introduzioni apportate in InceptionV3 prevedono la sostituzione di uno strato convoluzionale 5x5 in favore di due, meno espansivi, strati 3x3. Successivamente sono state applicate convoluzioni asimmetriche le quali, sostituendo convoluzioni 3x3 con convoluzioni 1x3 e 3x1, consentono di ridurre ulteriormente i costi computazionali. Per mantenere un alto grado di efficienza, l'architettura utilizza classificatori ausiliari come regolarizzatori per migliorare la convergenza delle reti neurali profonde e, di conseguenza, l'accuratezza nelle fasi successive. Infine, l'ultima modifica apportata, riguarda il ridimensionamento della griglia delle *feature maps* applicando due blocchi paralleli di convoluzione e *pooling* in seguito concatenati tra loro ottenendo, in tal modo, una griglia snodata.

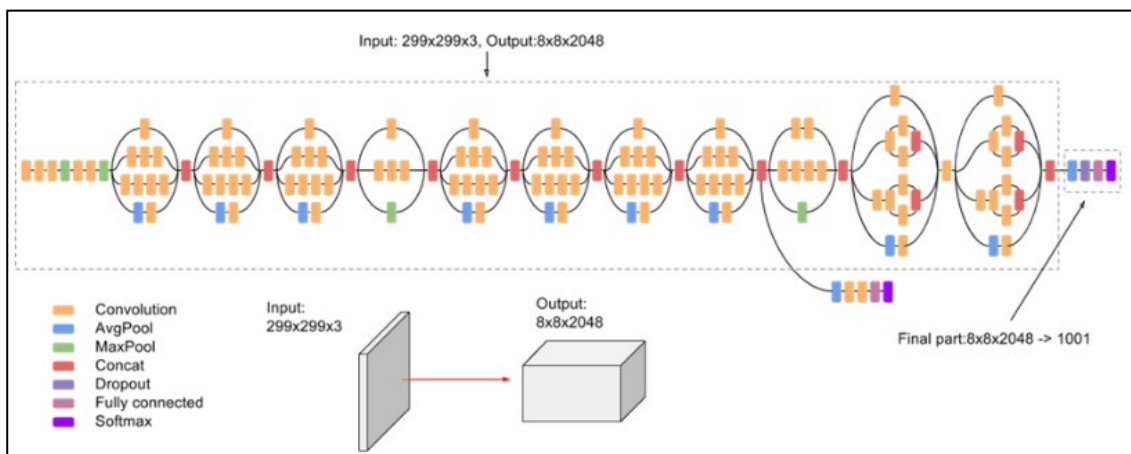


Figura 27: Architettura di InceptionV3²²⁶.

La finalità del task di *video activity classification* è la classificazione dei video prendendo in esame la sola parte visiva del documento. Tuttavia, per contestualizzare maggiormente il contenuto dei documenti audiovisivi, è necessario estrapolare e analizzare la componente sonora che costituisce, specie nei casi di documentari e interviste, una fonte inestimabile di conoscenza sul patrimonio orale come, ad esempio, termini linguistici e flessioni dialettali, testi di canti popolari o aneddoti folkloristici. Per l'estrazione di questa tipologia informativa esistono diverse librerie disponibili liberamente o a pagamento in Python. In questa sezione verranno passate in analisi alcune tra le più diffuse librerie evidenziandone le caratteristiche e differenze tra ciascuna di esse.

²²⁵ Christian Szegedy et al., "Rethinking the Inception Architecture for Computer Vision," (2015). <https://arxiv.org/abs/1512.00567>.

²²⁶ Google Cloud, "Guida avanzata a Inception v3," ultimo accesso 23 Gennaio 2024, <https://cloud.google.com/tpu/docs/inception-v3-advanced?hl=it>.

AssemblyAI²²⁷. È una API contenente diversi modelli di intelligenza artificiale per la trascrizione e la comprensione dei dati orali. Si tratta di una soluzione ad altissimo potenziale sviluppata, in particolare, per finalità lavorative e come tale disponibile a pagamento con la possibilità di uso gratuito mediante una demo²²⁸. In particolare, l'API si compone nei seguenti tre moduli:

1. *Transcribing*, costituisce il nucleo centrale dell'applicativo funzionale alla trascrizione dei dati orali restituendo in output un file strutturato in JSON.
2. *Summarizing*, costruito su tecniche di NLP per il trattamento del testo prodotto e l'estrazione delle informazioni salienti.
3. *Analyze*, modulo creato per finalità commerciali indirizzata a call center a estrapolare informazioni dai loro dati allo scopo di migliorare i propri servizi.

Pywit²²⁹. È un modello sviluppato da Wit.ai per la trascrizione dei dati orali e l'elaborazione del testo ottenuto. Similarmente ad AssemblyAi, anche Pywit è accessibile attraverso l'impiego di token ottenibili a pagamento. Le funzionalità di questo modello prevedono:

- **Message API**, finalizzata all'estrazione del significato da porzioni di testo
- **Speech API**, il modulo centrale per la trascrizione del parlato. I formati audio accettati per il processamento sono: WAV, MPEG3, OGG, ULAW e RAW
- **Interactive**, per avviare una conversazione con un proprio bot.

Google Cloud-Speech²³⁰. L'API di Google per il riconoscimento del parlato è tra le più sofisticate e avanzate al mondo e costituisce la base dei servizi di traduzione del colosso di Mountain View. È possibile usufruire a pagamento del servizio con le tecnologie aggiornate altrimenti è disponibile sfruttare una versione antecedente del servizio gratuitamente. Speech-to-Text impiega tre metodi principali per l'elaborazione del parlato²³¹:

- Riconoscimento sincrono. È un'operazione che accetta file audio della durata inferiore a un minuto. I risultati derivanti dall'elaborazione dei dati vengono restituiti dopo aver processato l'intero audio;
- Riconoscimento asincrono. Accetta file audio della durata massima di 480 minuti. È un'operazione a lunga esecuzione attraverso la quale i risultati possono essere ottenuti periodicamente durante il processo;

²²⁷ AssemblyAI, "Docs," ultimo accesso 23 Gennaio 2024, <https://www.assemblyai.com/docs/>.

²²⁸ La demo è vincolata al token associato all'account creato con una possibilità di utilizzo di 5h.

²²⁹ WIT-AI, "PyWit," ultimo accesso 23 Gennaio 2024, <https://github.com/wit-ai/pywit..>

²³⁰ Python, "Google-speech," ultimo accesso 23 Gennaio 2024, <https://pypi.org/project/google-speech/>.

²³¹ Google, "Speech-to-Text," ultimo accesso 23 Gennaio 2024, <https://cloud.google.com/speech-to-text/docs/basics?hl=it>.

- Riconoscimento dei flussi di dati. Esegue il riconoscimento dei dati audio forniti in tempo reale come l'acquisizione attraverso un microfono. In questa elaborazione, i dati vengono processati simultaneamente e i risultati possono essere ottenuti mentre il soggetto sta continuando a parlare.

Per fruire facilmente delle API gratuite di Google Speech, è possibile impiegare la libreria SpeechRecognition.

SpeechRecognition. Oltre a Google Speech, questa libreria supporta le API gratuite rilasciate da vari motori quali Wit, Microsoft Azure Speech, IBM Speech to text etc. In base alla fonte in input o alle operazioni attese, vengono attivati i diversi moduli di cui si compone la libreria. La tabella seguente fornisce un riepilogo della struttura di SpeechRecognition.

Il funzionamento del framework ha inizio con l'inserimento in input del file sonoro in formato WAV, AIFF AIFF-C e FLAC. Dopo aver selezionato il documento da trattare si può specificare quale dei tool contenuti impiegare per l'elaborazione e, nell'eventualità, specificare un segmento o una durata massima del documento da prendere in esame.

NOME	PROCESSO	ATTIVAZIONE
PyAudio	Elaborazione dell'audio proveniente da microfono	Microfono come sorgente in input
PocketSphinx-Python	Riconoscimento del parlato	Specificare come strumento di riconoscimento
Vosk	Riconoscimento del parlato	Specificare come strumento di riconoscimento
Google Cloud Speech Library for Python	Riconoscimento del parlato	Specificare come strumento di riconoscimento
FLAC	Codificare l'audio in ingresso per il processo di riconoscimento	Automatico
Whisper	Trascrizione dell'audio	Di default
Whisper API	Trascrizione dell'audio	Da specificare nel caso si vogliano utilizzare le API

		di OpenAI per questo applicativo
IBM	Riconoscimento del parlato	Specificare come strumento di riconoscimento
Microsoft Azure	Riconoscimento del parlato	Specificare come strumento di riconoscimento

Tabella 8: moduli presenti nella libreria SpeechRecognition.

5. Lo standard IIF per la valorizzazione della conoscenza estratta automaticamente tramite Reti Neurali Convoluzionali

L'obiettivo del presente capitolo è quello di porre alcune considerazioni sulla possibilità di applicare, congiuntamente, le reti neurali convoluzionali e lo standard IIF per l'estrazione e la formalizzazione della conoscenza dalle immagini del patrimonio culturale. Le considerazioni si fondano sulla possibilità di applicare efficacemente strumenti basati su CNNs per estrarre informazioni che possono costituire un valore aggiunto al patrimonio culturale e, di conseguenza, essere inclusi all'interno di un modello IIF per la disseminazione e condivisione delle risorse culturali. Come ulteriore punto di analisi, nel presente capitolo la trattazione è finalizzata a individuare gli aspetti mediante i quali il processo di capitalizzazione della conoscenza può essere effettuato automaticamente, distinguendo tre sub attività:

1. Informazioni estraibili mediante le reti neurali convoluzionali.
2. Metadati richiesti nel modello IIF.
3. Automatizzazione del processo.

Informazioni estraibili mediante le reti neurali convoluzionali

Una delle attività principali che denota l'efficacia di una rete neurale è l'addestramento, attraverso la quale viene impostato il target informativo del modello in oggetto come, ad esempio, riconoscere determinati oggetti posti in primo e secondo piano o parzialmente nascosti. In tale fase, è dunque essenziale disporre di un campione di addestramento (*training set*) chiaro ed esaustivo affinché le reti neurali possano produrre correttamente l'attività richiesta. Al fine di rispondere alle diverse esigenze provenienti dai vari domini, la comunità scientifica contribuisce ad alimentare la quantità di dataset, dal diverso grado di specificità, che gli utenti possono impiegare per le operazioni di addestramento o di verifica. In base alle modalità e al campione di addestramento impiegato, i risultati prodotti dalle reti neurali possono differenziarsi.

Nel dominio culturale, l'applicazione delle reti neurali convoluzionali è fortemente condizionato dall'elevato numero di elementi differenti tra loro. La varietà del patrimonio culturale rappresenta il punto centrale di riflessione attorno al quale occorre delineare il processo per un'applicazione delle reti neurali coerente con il dominio e la tipologia di bene culturale. La conoscenza di dominio è dunque imprescindibile poiché, altrimenti, sarebbe possibile individuare, a partire dalla fase preparatoria, le caratteristiche degli elementi da rintracciare e sulla base delle quali impostare la metodologia per la costruzione di un dataset specifico o l'adattamento di uno esistente.

In tale dominio, le reti neurali convoluzionali, permettendo di analizzare il contenuto di immagini e audiovisivi, possono essere applicate per una pluralità di finalità. Combinando tra loro tecniche di

riconoscimento e classificazione, ad esempio, è possibile individuare punti di deterioramento in oggetti o edifici consentendo il tempestivo avvio di attività di restauro e conservazione²³², rintracciare oggetti poco visibili all'interno di siti culturali²³³ o, ancora, rintracciare oggetti nei dipinti²³⁴ e riconoscerne lo stile artistico²³⁵. Nel caso di immagini contenenti elementi visivi e testuali, come ad esempio un manoscritto medievale, le reti neurali convoluzionali possono essere finalizzate a riconoscere i testi attraverso l'ICR e, al contempo, eventuali miniature e oggetti in esse contenuti. Per ciò che concerne la componente testuale, i sofisticati strumenti di ICR consentono un'estrazione altamente accurata di testi scritti a mano permettendone la trascrizione e la conversione in testo *machine readable*. Il riconoscimento degli oggetti, invece, permette di lavorare sulla parte visiva delle immagini e di estrarre elementi caratterizzanti unitamente alle coordinate spaziali della loro collocazione all'interno dell'immagine.

Con esplicito riferimento agli audiovisivi, le tecniche di *activity recognition* basate su reti neurali convoluzionali permettono di riconoscere le attività manifestate all'interno mediante la classificazione dei movimenti. Il riconoscimento delle attività può assumere un ruolo di rilievo nel contesto del patrimonio culturale immateriale se finalizzato a identificare espressioni delle tradizioni folkloristiche e antropologiche quali danze popolari, processioni e attività di vario genere (es. alpinismo, artigianato etc.) ognuna delle quali con proprie caratteristiche e movimenti.

Beneficiando dell'alta prestazione e velocità di calcolo delle reti neurali convoluzionali, è dunque possibile rintracciare, da immagini di ogni formato, potenziali elementi del patrimonio culturale in grado di arricchire la base di conoscenza per una descrizione standardizzata delle risorse culturali attraverso lo standard IIF. Come ulteriore elemento di raffinamento semantico, gli elementi risultati dall'applicazione delle CNNs possono essere formalizzati ricorrendo al modello ontologico CIDOC Conceptual Reference Model (CRM), promosso dall'International Committee for Documentation (CIDOC) dell'International Council of Museums (ICOM) per la rappresentazione delle informazioni e degli oggetti del dominio culturale. Seppure l'obiettivo primario del modello ontologico sia quello di definire lo scambio e la mediazione di informazioni strutturate, è possibile formalizzare anche risorse non strutturate senza che queste abbiano una forma strutturata o logica. Nell'attuale

²³² Rachna Pathak, Anil Saini, Arnav Wadhwa, Himanshu Sharma, and Dhiraj Sangwan, "An object detection approach for detecting damages in heritage sites using 3-D point clouds and 2-D visual data," *Journal of Cultural Heritage* 48, (2021). DOI: 10.1016/j.culher.2021.01.002.

²³³ Wouter Verschoof-van der Vaart, and Lambers, Karsten, "Learning to Look at LiDAR: The Use of R-CNN in the Automated Detection of Archaeological Objects," *LiDAR Data from the Netherlands*, 2 (2019) DOI: 10.5334/jcaa.32.

²³⁴ Maria-Cristina Marinescu, Artem Reshetnikov, and Joaquim More López, "Improving object detection in paintings based on time contexts," *International Conference on Data Mining Workshops (ICDMW)*, (2020). DOI: 10.1109/ICDMW51313.2020.00133.

²³⁵ Saqib Imran, Rizwan Ali Naqvi, Muhammad Sajid, Tauqeer Safdar Malik, Saif Ullah, Syed Atif Moqurrab, and Dong Keon Yon, "Artistic Style Recognition: Combining Deep and Shallow Neural Networks for Painting Classification," *Mathematics* 11, n. 22 (2023). DOI: 10.3390/math11224564.

versione²³⁶, il modello CIDOC CRM si compone di 81 classi, «*general notions in the domain of discourse*»²³⁷, e 161 proprietà, «*the binary relations that link the individuals in the domain of discourse*»²³⁸, che consentono di descrivere il dominio culturale e «*the exchange and integration of information from heterogeneous sources for the reconstruction and interpretation of the past at a human scale, based on all kinds of material evidence, including texts, audio-visual material and oral tradition*»²³⁹.

Nel contesto proposto in tale lavoro di tesi, le classi e proprietà del modello ontologico permettono di attribuire alle informazioni estraibili dalle immagini una valenza semantica. A tal fine, occorre individuare le classi e proprietà che meglio possano rappresentare le informazioni estraibili. In linea con quanto recuperabile mediante le CNNs, è possibile distinguere:

- Testo. Dai testi estratti attraverso ICR è possibile applicare tecniche di NLP per recuperare entità che possono essere mappate automaticamente nel modello ontologico²⁴⁰.
- Classi di oggetti. Gli elementi estraibili attraverso l'*object detection* possono essere vari e comprendere differenti classi del modello ontologico come, ad esempio: *E19 Physical Object* «*items of a material nature that are units for documentation and have physical boundaries that separate them completely in an objective way from other objects*»²⁴¹, *E20 Biological Object* che, al contrario dell'oggetto fisico «*individual items of a material nature, which live, have lived or are natural products of or from living organisms*»²⁴².
- Classi di attività. I risultati derivanti dall'operazione di *activity recognition* possono essere formalizzati all'interno del CIDOC CRM attraverso la classe *E7 Activity* che comprende «*actions intentionally carried out by instances of E39 Actor that result in changes of state in the cultural, social, or physical systems documented*»²⁴³.

Metadati del modello IIF

²³⁶ Allo stato attuale, il modello ontologico è edito nella versione consolidata 7.1.2 e in un draft alla versione 7.2.3. Cfr. International Committee for Documentation (CIDOC) Conceptual Reference Model (CRM), "Versions of CIDOC-CRM," ultimo accesso 27 Gennaio 2024, <https://www.cidoc-crm.org/versions-of-the-cidoc-crm>.

²³⁷ International Committee for Documentation (CIDOC), and International Council of Museums (ICOM), *Conceptual Reference Model (CRM) v. 7.1.2* (2022), p. 20, https://www.cidoc-crm.org/sites/default/files/cidoc_crm_v7.1.2.pdf

²³⁸ International Committee for Documentation (CIDOC), and International Council of Museums (ICOM), *Conceptual Reference Model (CRM) v. 7.1.2 cit.*, p. 20, https://www.cidoc-crm.org/sites/default/files/cidoc_crm_v7.1.2.pdf

²³⁹ International Committee for Documentation (CIDOC), and International Council of Museums (ICOM), *Conceptual Reference Model (CRM) v. 7.1.2 cit.*, p. 9, https://www.cidoc-crm.org/sites/default/files/cidoc_crm_v7.1.2.pdf.

²⁴⁰ Dora Melo, Irene Pimenta Rodrigues, and Davide Varagnolo, "A strategy for archives metadata representation on CIDOC-CRM and knowledge discovery," *Semantic Web* 14, (2023). DOI: 10.3233/SW-222798.

²⁴¹ International Committee for Documentation (CIDOC), and International Council of Museums (ICOM), *Conceptual Reference Model (CRM) v. 7.1.2 cit.*, pp. 72-73, https://www.cidoc-crm.org/sites/default/files/cidoc_crm_v7.1.2.pdf.

²⁴² International Committee for Documentation (CIDOC), and International Council of Museums (ICOM), *Conceptual Reference Model (CRM) v. 7.1.2 cit.*, p. 73, https://www.cidoc-crm.org/sites/default/files/cidoc_crm_v7.1.2.pdf.

²⁴³ International Committee for Documentation (CIDOC), and International Council of Museums (ICOM), *Conceptual Reference Model (CRM) v. 7.1.2 cit.*, p. 65, https://www.cidoc-crm.org/sites/default/files/cidoc_crm_v7.1.2.pdf.

Lo standard IIIF costituisce, al giorno d'oggi, uno dei principali veicoli per la condivisione di materiale culturale digitale (manoscritti, immagini, dipinti, oggetti 3D etc.). Nell'ecosistema digitale, il IIIF è stato oggetto di innovazioni sempre più considerevoli che lo hanno portato ad essere la principale soluzione adottata di importanti istituti culturali. Per gli obiettivi posti in tale lavoro, la Presentation API v. 3 costituisce la scelta migliore per la definizione del modello di Manifest IIIF potendo includere, oltre a immagini, anche audiovisivi. La Presentation API, la cui struttura è stata oggetto di trattazione nel secondo capitolo, definisce un set di metadati necessari, quindi obbligatori, per la corretta implementazione del modello mentre altri, opzionali, possono essere impiegati per arricchire la descrizione degli oggetti rappresentati. Tra questi, l'elemento `<metadata>` è implementato per offrire agli utenti la possibilità di aggiungere liste di metadati al di fuori del IIIF (Figura 28).

```
{
  "metadata": [
    {
      "label": { "en": [ "Creator" ] },
      "value": { "en": [ "Anne Artist (1776-1824)" ] }
    }
  ]
}
```

Figura 28: elemento "metadata" all'interno delle Presentation API 3.0²⁴⁴.

Per la corretta valorizzazione della proprietà `<metadata>` gli utenti devono rispettare la sintassi JSON-LD *label-value* senza, tuttavia, essere vincolati all'uso di specifici standard di metadati quali, per citarne alcuni, Dublin Core²⁴⁵ o VRA Core²⁴⁶. L'assenza di vincoli legati all'adozione di uno standard specifico di metadati, permette di configurare la proprietà `<metadata>` per effettuare l'ingest delle informazioni estraibili tramite le reti neurali convoluzionali nelle classi CIDOC CRM incrementando il livello descrittivo della risorsa all'interno del modello IIIF.

In aggiunta a ciò, gli elementi estraibili attraverso le reti neurali possono essere rappresentati attraverso il sistema di annotazione, una delle funzionalità di maggior popolarità all'interno dello standard che permette di evidenziare, secondo le specifiche del visualizzatore applicato, gli elementi contenuti nell'immagine e ad associarvi ulteriori informazioni come link esterni, testo etc. Le coordinate della posizione degli elementi all'interno del file sono uno dei requisiti necessari per la

²⁴⁴ International Image Interoperability Framework, "Presentation API," ultimo accesso 27 Gennaio 2024, <https://iiif.io/api/presentation/3.0/#31-descriptive-properties>

²⁴⁵ ISO 15836-2003 *Information and documentation - The Dublin Core™ metadata element set* (2003), <https://www.dublincore.org/specifications/dublin-core/dcmi-terms/>.

²⁴⁶ Library of Congress, *Visual Resources Association Core (VRA Core) v.4* (2007), <https://www.loc.gov/standards/vracore/schemas.html>.

loro corretta visualizzazione mediante il sistema di annotazione. In fase di implementazione, il sistema di annotazione è stato caratterizzato da finalità legate allo studio e all'analisi dei manoscritti, per fornire agli utenti uno strumento attraverso cui elaborare commenti, trascrizioni o semplici appunti. Tramite un apposito *annotation server*, l'utente che vi accede ha modo di visualizzare i contenuti proposti dall'autore del modello IIIF e, al contempo, di effettuare proprie annotazioni. Allo stato attuale, con sempre un maggior numero di oggetti rappresentabili attraverso lo standard, la funzionalità di annotazione è impiegata, da utenti e creatori di contenuti, per evidenziare nelle immagini diversi dettagli di interesse. Per le finalità poste in essere in tale capitolo, il sistema di annotazione può essere impiegato per dare un riscontro visivo alle informazioni estraibili attraverso le reti neurali, distinguendo tra:

- Testi: la parte testuale delle immagini, estraibile tramite sistemi di riconoscimento automatico della scrittura, può essere usata all'interno del sistema di annotazione come mezzo di confronto con il testo presente nell'immagine.
- Classi di oggetti: gli elementi risultati dalle attività di *object detection*, insieme alle proprie coordinate spaziali della posizione all'interno dell'immagine, possono essere direttamente strutturati nel Manifest IIIF tramite il sistema di annotazioni.
- Classi di attività: similmente alle classi di oggetti, anche le informazioni estraibili con l'operazione di *activity recognition* possono essere veicolate mediante il sistema di annotazione. Nel caso degli audiovisivi, occorre specificare le coordinate temporali unitamente a quelle spaziali per specificare gli esatti momenti in cui le azioni hanno luogo.

Automatizzazione del processo

Dopo aver stabilito i punti di contatto per mezzo dei quali le informazioni estraibili tramite le reti neurali convoluzionali possono essere capitalizzate all'interno dello standard IIIF, occorre individuare le modalità attraverso cui il processo possa avvenire automaticamente. Tale attività richiede, preliminarmente, l'impiego di strumenti conformi alla Presentation API adottata per la generazione dei Manifest IIIF. L'attuale scenario presenta un cospicuo numero di implementazioni, basate su differenti tecnologie, che guidano gli utenti nella generazione dei modelli IIIF. Nella maggior parte dei casi, si tratta di strumenti con un'interfaccia *user friendly* che, tuttavia, richiedono l'intervento umano per la compilazione dei metadati. Solamente pochi strumenti mostrano le potenzialità per automatizzare il processo come nei casi della libreria python *prezi3*²⁴⁷ e Science

²⁴⁷ Mike Bennet, and Giacomo Marchioro, *iiif-prezi3 v. 1.2.1* (2023). <https://github.com/iiif-prezi/iiif-prezi3>.

Stories API²⁴⁸, un applicativo basato su diverse tecnologie web, sviluppato nel contesto Wikidata per la generazione automatica dei Manifest a partire dalle immagini contenute nel repository.

Per le finalità descritte, l'impiego di *prezi3* costituisce la soluzione più adatta per importare le informazioni estraibili all'interno del modello IIF. Il criterio che fa propendere verso la scelta della libreria, è la possibilità di configurare all'interno di un unico ambiente, python, l'applicazione delle reti neurali e la successiva generazione dei Manifest IIF, consentendo di stabilire a priori delle regole di associazione attraverso cui mappare nel modello gli elementi estraibili secondo le classi CIDOC CRM.

Sulla base alle considerazioni poste all'interno del presente capitolo, è possibile delineare schematicamente il workflow che, partendo dall'applicazione di reti neurali convoluzionali, mira a formalizzare all'interno dello standard IIF la conoscenza estraibile da immagini del patrimonio culturale (Figura 29). L'esistenza di librerie, modelli di CNNs e strumenti testati e largamente impiegati, consente di gestire l'intero processo in python.

Come anticipato, la fase iniziale concerne l'estrazione del contenuto informativo tramite le reti neurali convoluzionali. L'applicazione delle CNNs può avvenire per processare immagini e audiovideo i cui risultati, per la maggior parte testi, possono essere descritti mediante il modello ontologico CIDOC CRM ricorrendo a librerie python esistenti quale, ad esempio, *cidoc-crm-neo4j*²⁴⁹. Tale operazione richiede l'identificazione delle classi del modello ontologico da richiamare, all'interno della libreria, per dare valenza semantica agli elementi ottenibili tramite le CNNs.

La fase successiva ha per oggetto l'ingest delle informazioni all'interno del modello IIF mediante una duplice attività: la prima è finalizzata a importare le descrizioni all'interno della proprietà *<metadata>*, mentre la seconda ha l'obiettivo di valorizzare le informazioni sotto l'aspetto visivo ricorrendo al sistema di annotazione. La libreria *prezi3* permette di definire la struttura del modello di Manifest IIF che, attraverso regole di associazione, può essere valorizzata con le informazioni risultati dall'applicazione delle reti neurali convoluzionali.

²⁴⁸ Katherine Thornton, and Kenneth Seals-Nutt, "Science Stories: Using IIF and Wikidata to Create a Linked-Data Application," *International Workshop on the Semantic Web 2180*, (2018). <https://ceur-ws.org/Vol-2180/paper-68.pdf>.

²⁴⁹ The Digital Innovation Group, Arizona State University, "cidoc-crm-neo4j," ultimo accesso 27 Gennaio 2024, <https://github.com/diging/cidoc-crm-neo4j>.

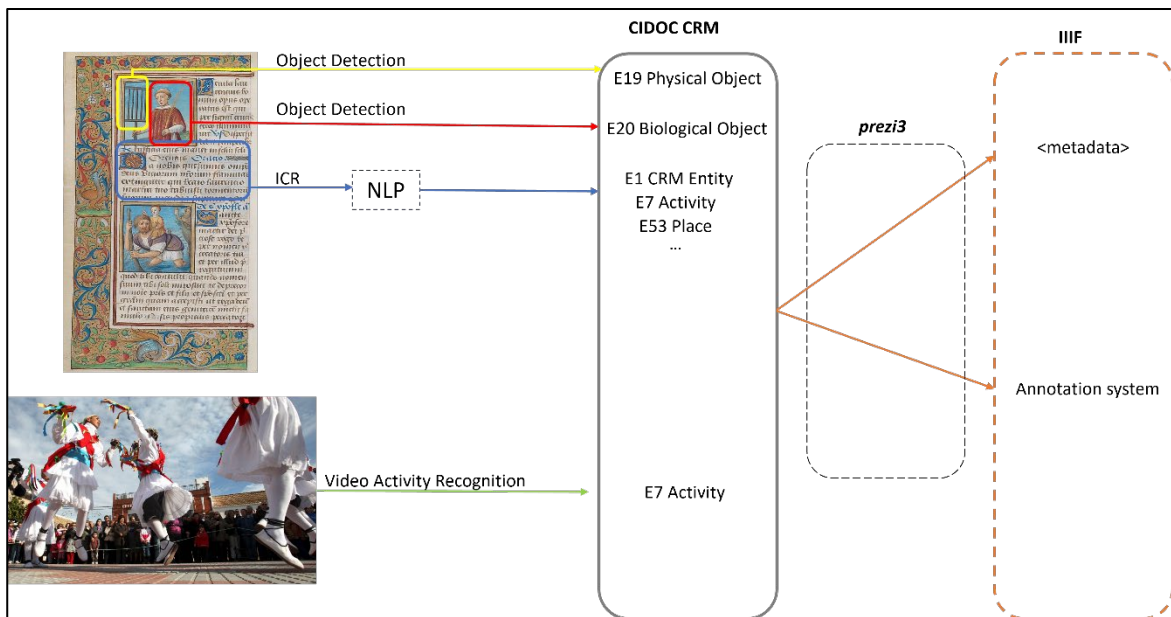


Figura 29: Rappresentazione del workflow di estrazione automatica delle informazioni tramite reti neurali e formalizzazione attraverso lo standard IIF.

Il risultato del processo sopra descritto è una rappresentazione in IIF di immagini del patrimonio culturale, ulteriormente valorizzate degli elementi ottenibili mediante l'applicazione di strumenti di estrazione automatica della conoscenza basati su reti neurali convoluzionali. Gli aspetti che connotano il processo riguardano la possibilità di formalizzare gli elementi con il modello ontologico CIDOC CRM e l'automatizzazione del processo per importare gli elementi nel modello di Manifest IIF. L'espedito adottato in Wikidata, attraverso lo sviluppo del citato software Science Stories API, dimostra, per le immagini presenti nel repository, la possibilità di creare automaticamente Manifest IIF valorizzabili con le informazioni provenienti da Wikipedia.

Il modello di Manifest IIF arricchito con le informazioni estraibili automaticamente, oltre a rappresentare una straordinaria modalità di disseminazione di risorse digitali del patrimonio culturale, costituisce una potenziale fonte informativa che può essere acquisita da repository ed altre banche dati culturali. Nel caso di Europeana, ad esempio, l'acquisizione delle informazioni dai Manifest IIF favorisce la promozione di contenuti standardizzati e *open access*. A tale scopo, è stato sviluppato un framework in grado di mappare il protocollo IIF con l'*Europeana Data Modelling (EDM)*²⁵⁰ per promuovere gli oggetti culturali digitali attraverso l'*Europeana Cloud API* cosicché da contribuire «to an open, knowledgeable and creative society»²⁵¹. L'impiego di strumenti automatici come il

²⁵⁰ Europeana Cloud, *D4.4 Recommendations for enhancing EDM to represent digital Content* (2016),

https://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/Europeana_Cloud/Deliverables/D4.4%20Recommendations%20For%20Enhancing%20EDM%20to%20Support%20Research%20Oriented%20Content.pdf

²⁵¹ Antoine Isaac, "The state of IIF at Europeana," *OCLC Research Mini-symposium on the Discovery and Use of Open Collections*, (2019), p. 8, <https://www.oclc.org/content/dam/research/presentations/2019/061919-Isaac.pdf>.

*Digital Aggregation Lab Software*²⁵² consente di automatizzare il processo di acquisizione delle informazioni dai Manifest IIF tramite la definizione di un workflow che mira a individuare ed estrarre i metadati di interesse per l'EDM.

Le considerazioni poste nel presente capitolo hanno lo scopo di mettere in luce i vantaggi derivabili dall'applicazione congiunta delle tecnologie e standard, analizzati nel lavoro di tesi, per la valorizzazione del patrimonio culturale. In particolare, il processo qui delineato prende in esame le modalità attraverso cui gli elementi estraibili da immagini, o audiovideo, del patrimonio culturale possano essere automaticamente acquisiti all'interno di un modello di Manifest IIF per arricchire la descrizione degli oggetti digitali. Inoltre, l'adozione di un modello ontologico specifico del patrimonio culturale come il CIDOC CRM contribuisce ad impreziosire la descrizione sotto l'aspetto semantico.

La valorizzazione delle risorse culturali sotto l'aspetto descrittivo costituisce un primo passo verso la promozione del patrimonio culturale che, attraverso la rappresentazione in IIF, può essere condiviso a beneficio di tutti gli attori coinvolti nel settore culturale. In primo luogo, gli istituti culturali potranno giovare di un framework in grado di dare valore alle proprie collezioni in chiave digitale e, al contempo, attirare un maggiore interesse verso il patrimonio culturale. Insieme agli istituti, anche i *player* di mercato saranno messi nella condizione di offrire servizi aggiuntivi e, di conseguenza, concorrere ad elevare la qualità dei prodotti culturali digitalizzati di cui potranno beneficiare utenti, studiosi, ricercatori o generici interessati, e i territori.

6. Prospettive future e conclusioni.

L'importanza della standardizzazione nell'ecosistema digitale

Il digitale sta progressivamente ridefinendo l'epoca attuale e trasformando il mondo in cui viviamo mediante una incessante produzione di dati perpetrata attraverso le più svariate modalità come, ad esempio, l'invio di e-mail, l'acquisizione di fotografie tramite smartphone o altro dispositivo, le transazioni di acquisto su siti *e-commerce* etc. L'attuale società è dunque caratterizzata dalla trasformazione al digitale che, tramite la reingegnerizzazione dei processi, sta progressivamente soppiantando l'analogico. La pandemia da COVID-19 ha sicuramente avuto un ruolo non secondario nella transizione, forzando cittadini, imprese e Governi verso servizi e attività digitali (Figura 30).

²⁵² Nuno Freire, "The Data Aggregation Lab Software. Experimentation for Linked Data Aggregation in Cultural Heritage," *Metadata and Semantic Research. MTSR 2019. Communications in Computer and Information Science* 1057, (2019). DOI: 10.1007/978-3-030-36599-8_38.

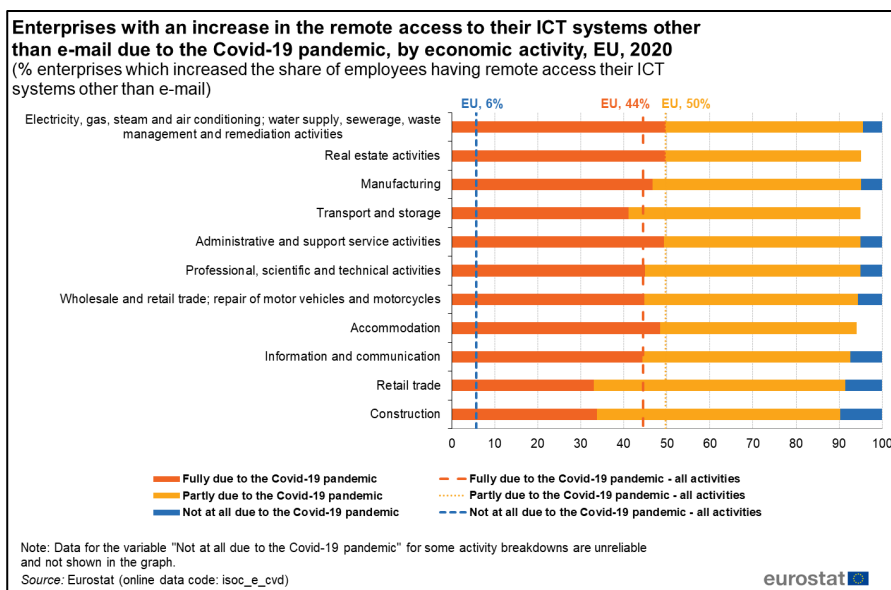


Figura 30: Aziende con un incremento negli accessi in remoto ai propri servizi ICT fatto salvo le email a causa della pandemia da Covid-19²⁵³.

Il grafico nella Figura 28 mostra i risultati dell'indagine condotta da Eurostat in merito al legame COVID-19 e incremento dei servizi ICT promossi dalle aziende. I dati hanno evidenziato, prendendo ad esame la media dei settori economici, la connessione tra i due fattori totalizzando tra la piena (44%) e la parziale attribuzione (50%) oltre il 90% di corrispondenza. Nell'attuale ecosistema digitale la produzione di Big Data, espressione comunemente impiegata per indicare l'insieme di dati «vasti e complessi»²⁵⁴ raccolti tramite fonti diverse, costituiscono una ricchezza informativa dalla quale i *player* dei mercati economici, e in particolare i giganti dell'informatica, attingono per realizzare i propri interessi. L'incerta, o talvolta assente, definizione di procedure e metodologie standardizzate comporta, a livello internazionale, l'assunzione di pratiche soggettive che contribuiscono a incrementare la frammentazione e l'isolamento di processi e attività che, al contrario, dovrebbero poter comunicare tra loro. Nel contesto digitale, fortemente permeato dalle applicazioni delle *Information and Communication Technologies (ICT)*, la standardizzazione garantisce livelli minimi di sicurezza e affidabilità dei sistemi rafforzando, di conseguenza, la collaborazione tra le diverse comunità.

Nel dominio culturale, e in particolare nei processi di digitalizzazione, la disposizione di tecnologie, processi e strumenti standardizzati è un requisito fondamentale per la realizzazione attività di tutela

²⁵³ Eurostat, "Enterprises with an increase in the remote access to their ICT systems due to the Covid-19, by activity, EU 2020," ultimo accesso 23 Gennaio 2024, https://ec.europa.eu/eurostat/statistics-explained/index.php?title=File:Enterprises_with_an_increase_in_the_remote_access_to_their_ict_systems_due_to_the_Covid-19_by_activity_EU_2020.png.

²⁵⁴ Parlamento Europeo, "Big data: definizione, benefici e sfide," ultimo accesso 26 Gennaio 2024, <https://www.europarl.europa.eu/news/it/headlines/society/20210211STO97614/big-data-definizione-benefici-e-sfide-infografica>.

e valorizzazione del patrimonio culturale. Il Piano nazionale di digitalizzazione del patrimonio culturale testimonia tale necessità indicando nelle Linee guida per i processi di digitalizzazione 2022-2023²⁵⁵ standard di riferimento per le diverse attività prese in esame. Ad esempio:

- **Metadati:** indicando standard di metadati descrittivi specifici del settore di riferimento delle risorse digitali come ISAD (G) (*General International Standard Archival Description*)²⁵⁶ ed EAD (*Encoded Archival Description*)²⁵⁷ per il dominio archivistico o le IFLA-ISBD (*International Federation of Library Associations and Institutions-International Standard Bibliographic Description*)²⁵⁸ per quello bibliografico, o il METS (*Metadata Encoding and Transmission Standard*)²⁵⁹ come standard per la codifica e l'interscambio dei metadati tra i sistemi. Sono altresì indicati standard di metadati tecnici in base al formato della risorsa e standard amministrativo-gestionale per indicare diritti e licenze connesse alle risorse insieme e fornire dettagli utili alla conservazione delle risorse. In quest'ultimo caso lo standard preso a riferimento è il PREMIS (*Preservation Metadata: Implementation Strategies*)²⁶⁰, appositamente definito per la conservazione permanente di oggetti digitali.
- **Strumenti:** per ciascuna attrezzatura necessaria nei processi di digitalizzazione sono indicate le caratteristiche e le impostazioni che queste devono avere. Nel caso delle digitalizzazioni 2D, ad esempio, sono dettagliate le tipologie di scanner da impiegare (piani, planetari, a tamburo etc.) e la risoluzione minima da impostare per ciascuno di essi. Similarmente, anche per altre attività come le scansioni 3D, le fotografie o l'acquisizione di audiovisivi, sono indicati i requisiti e le impostazioni da adottare per ottenere un lavoro qualitativamente conforme alle Linee guida²⁶¹.
- **Formati:** la scelta del formato incide fortemente sulla possibilità di conservazione e di futuro accesso della risorsa digitalizzata. Le Linee guida forniscono le indicazioni

²⁵⁵ Istituto Centrale per la Digitalizzazione del Patrimonio Culturale, *Linee guida per i processi di digitalizzazione 2022-2023* (2022), <https://docs.italia.it/italia/icdp/icdp-pnd-digitalizzazione-docs/it/v1.0-giugno-2022/index.html>.

²⁵⁶ International Council on Archives, *ISAD (G): General International Standards Archival Description – Second edition* (2000), <https://www.ica.org/en/isadg-general-international-standard-archival-description-second-edition>.

²⁵⁷ Society of American Archivists, *Encoded Archival Description (EAD3)* (2019), <https://www.ica.org/en/isadg-general-international-standard-archival-description-second-edition>.

²⁵⁸ International Federation of Library Associations and Institutions, *ISBD: International Standard Bibliographic Description. Consolidated edition*. Edizione italiana a cura dell'ICCU, 2012. <https://www.iccu.sbn.it/it/eventi-novita/novita/ISBD-International-Standard-Bibliographic-Description.-Consolidated-edition.-Edizione-italiana-novembre-2012/>.

²⁵⁹ Library of Congress, *Metadata Encoding & Transmission Standard* v. 1.12.1, <https://www.loc.gov/standards/mets/mets-schemadocs.html>

²⁶⁰ Library of Congress, *PREMIS Data Dictionary for Preservation Metadata* v.3.0 (2015), <https://www.loc.gov/standards/premis/v3/premis-3-0-final.pdf>.

²⁶¹ Cfr. Istituto Centrale per la Digitalizzazione del Patrimonio Culturale, *Linee guida per i processi di digitalizzazione 2022-2023 cit.* Cap. 2.4 "Come," <https://docs.italia.it/italia/icdp/icdp-pnd-digitalizzazione-docs/it/v1.0-giugno-2022/il-progetto-di-digitalizzazione/come.html>.

necessarie per guidare gli utenti verso l'adozione di formati di file aperti²⁶² e standard²⁶³. A tale scopo è richiamato l'Allegato 2 – Formati di file e riversamento²⁶⁴ delle Linee guida AGID sulla formazione, gestione e conservazione dei documenti informatici²⁶⁵. L'allegato analizza un numero ristretto di formati²⁶⁶ valutandone, in base alle caratteristiche, l'adozione per la lettura, scrittura e conservazione delle specifiche tipologie di documenti informatici (Figura 31).

XML	FORMATO DI FILE
Nome completo	Extensible Markup Language
Estensione/i	.xml
Magic number	<?xml 0x20
Tipo MIME	application/xml, text/xml
Sviluppato da	World Wide Web Consortium
Tipologia di standard	aperto, estendibile, retrocompatibile, <i>de iure</i> , testuale
Livello metadati	4
Derivato da	SGML
Revisione	1.0, 5ª edizione
Riferimenti	<ul style="list-style-type: none"> • W3C Recommendation XML 1.0 (5th Ed.), 2013 • www.w3.org/standards/xml • validator.w3.org
Conservazione	Sì, se conservato insieme a un XML Schema
Racc. per la lettura	Generico; subordinato ad eventuali, ulteriori obblighi o raccomandazioni di conformità con schemi/dialetti.
Racc. per la scrittura	Generico; subordinato ad eventuali, ulteriori obblighi o raccomandazioni di conformità con schemi/dialetti.

Figura 31: Scheda di valutazione del formato XML per gli ipertesti²⁶⁷.

Unitamente agli elementi appena descritti, le Linee guida per i processi di digitalizzazione 2022-2023 rimandano a specifici standard, nazionali e internazionali, che definiscono metodologicamente i

²⁶² Un formato è detto aperto o pubblico quando «è descritto da specifiche pubbliche, liberamente accessibili [...] ed esaustive». Cfr. Stefano Allegrezza, “I requisiti per i formati elettronici” in *La produzione e conservazione del documento digitale*, ed. Stefano Pigliapoco, and Stefano Allegrezza (Macerata: EUM, 2008), p. 157. ISBN: 978-88-6056-158-9.

²⁶³ Agenzia per l'Italia Digitale (AGID), *Linee guida per la formazione, gestione e conservazione dei documenti digitali* (2021), https://www.agid.gov.it/sites/default/files/repository_files/linee_guida_sul_documento_informatico.pdf.

²⁶⁴ Agenzia per l'Italia Digitale (AGID), Allegato 2 “Formati di file e riversamento” in *Linee guida per la formazione, gestione e conservazione dei documenti digitali* (2021), https://www.agid.gov.it/sites/default/files/repository_files/allegato_2_formati_di_file_e_riversamento.pdf.

²⁶⁵ In merito alla standardizzazione si distingue fra standard *de iure*, ossia ufficialmente definito da un ente di standardizzazione come, ad esempio, ISO (International Organization for Standardization) a livello internazionale o UNI (Ente Nazionale di Unificazione) per il contesto italiano, e standard *de facto*, ossia un formato largamente utilizzato dalle comunità di utenti tale da divenire popolarmente adottato senza che le sue specifiche siano state definite da un ente di standardizzazione. Nella valutazione la preferenza è rivolta verso gli standard riconosciuti “di diritto” che, come tali, godono di un riconoscimento accreditato della qualità delle proprie specifiche. Cfr. Stefano Allegrezza, “I requisiti per i formati elettronici” *cit.*, pp. 160-161. ISBN: 978-88-6056-158-9.

²⁶⁶ Sono presi in considerazione i formati tra quelli che «possono maggiormente garantire il principio di interoperabilità tra i sistemi di gestione documentale e conservazione e in base alla normativa vigente riguardante specifiche tipologie di documenti». Cfr. Agenzia per l'Italia Digitale (AGID), Allegato 2 “Formati di file e riversamento” *cit.*, p. 3. https://www.agid.gov.it/sites/default/files/repository_files/allegato_2_formati_di_file_e_riversamento.pdf.

²⁶⁷ Cfr. Agenzia per l'Italia Digitale (AGID), Allegato 2 “Formati di file e riversamento” *cit.*, p. 34. https://www.agid.gov.it/sites/default/files/repository_files/allegato_2_formati_di_file_e_riversamento.pdf

processi connessi con la digitalizzazione delle risorse culturali. Per dare alcuni esempi, tra questi rientrano:

- *Guidelines on the Production and Preservation of Digital Audio Objects, IASA-TC04, Second Edition*²⁶⁸.
- Linee guida per pianificare la digitalizzazione di collezioni di libri rari e manoscritti²⁶⁹.
- *Basic principles and tips for 3D digitisation of tangible cultural heritage for cultural heritage professionals and institutions and other custodians of cultural heritage*²⁷⁰.

A partire dunque dalla lettura del PND e delle Linee guida si evince la propedeuticità della standardizzazione alla realizzazione di attività indirizzate alla promozione del patrimonio culturale. Ciò si manifesta con maggior chiarezza laddove, ad essere applicate, sono le tecnologie di intelligenza artificiale, tanto potenti quanto rischiose che, come tali, necessitano di essere regolarizzate.

Verso una standardizzazione dell'intelligenza artificiale

In concomitanza alla crescente proliferazione delle tecnologie e del conseguente uso da parte di un numero sempre maggiore di individui, si avverte la necessità di avviare percorsi volti a standardizzare e rendere “sicura” l’AI. Si tratta di un tema tanto dibattuto a livello europeo e internazionale in cui principalmente ci si interroga su quali debbano essere gli aspetti regolamentati e il ruolo e la responsabilità di programmatori, sviluppatori e *stakeholders* nell’uso dell’intelligenza artificiale. Il tema è estremamente complesso sul piano giuridico dove occorre accertare la colpa del produttore, in ambito civile e penale, diversamente dalle finalità e modalità d’uso di singoli utenti. In tal senso, è necessario appurare fino a che punto la responsabilità è da rintracciarsi nel produttore, tenendo conto dell’imprevedibilità del comportamento di alcuni dispositivi intelligenti²⁷¹, e dove subentra invece quella del consumatore che impiega tali sistemi per perseguire propri fini. La realizzazione di una intelligenza artificiale sicura e affidabile è un processo che coinvolge l’ausilio di diversi attori, a livello nazionale e internazionale, chiamati a definire politiche, buone pratiche e norme a riguardo. È

²⁶⁸ International Association of Sound and Audiovisual Archives (IASA), *Guidelines on the Production and Preservation of Digital Audio Objects, IASA-TC04, Second Edition* (2009), <https://www.iasa-web.org/tc04/audio-preservation>.

²⁶⁹ International Federation of Library Associations and Institutions - Rare Book and Special Collections Section, *Linee guida per pianificare la digitalizzazione di collezioni di libri rari e manoscritti* (2014), <https://repository.ifla.org/bitstream/123456789/458/1/guidelines-for-planning-digitization-it.pdf>.

²⁷⁰ European Commission, Expert Group on Digital Cultural Heritage and Europeana, *Basic principles and tips for 3D digitisation of tangible cultural heritage for cultural heritage professionals and institutions and other custodians of cultural heritage* (2020), <https://digital-strategy.ec.europa.eu/en/library/basic-principles-and-tips-3d-digitisation-cultural-heritage>.

²⁷¹ Beatrice Fragrasso, “La responsabilità penale del produttore di sistemi di intelligenza artificiale,” *Diritto Penale Contemporaneo. Rivista Trimestrale*, 1/2023, (2023), https://www.sistemapenale.it/pdf_contenuti/1686584973_fragasso-la-responsabilita-penale-del-produttore-di-sistemi-di-ia.pdf.

stato questo uno dei temi di maggiore discussione nel corso del Forum economico mondiale²⁷², durante il quale alcuni dei protagonisti dell'intelligenza artificiale hanno sottolineato tale esigenza in ragione di quanto avvenuto con il «far west dei social media»²⁷³ posti per anni al di fuori di ogni regolarizzazione compromettendo la riservatezza dei dati degli utenti. L'esempio più eclatante è, con ogni probabilità, quanto avvenuto con Cambridge Analytica, la società rea di aver ottenuto con l'inganno i dati di oltre 50 milioni di utenti Facebook per condizionare la campagna elettorale di Trump alla Presidenza degli Stati Uniti nel 2016²⁷⁴. Le preoccupazioni che un tale scenario si ripresenti con l'intelligenza artificiale sono forti e in parte giustificate, riprendendo le parole di Sam Altman, co-fondatore di OpenAI, al Forum «C'è una parte delle preoccupazioni che è corretta: questa tecnologia è molto, molto potente e non sappiamo cosa può succedere. Può andare in modo molto sbagliato. Ma possiamo prendere precauzioni per renderla sicura»²⁷⁵.

Inserita in tale contesto, la standardizzazione assume dunque una nuova dimensione, interessando l'intero ecosistema digitale e non singoli elementi. La necessità di attuare una normalizzazione di tali tecnologie è quantomai impellente per un contesto, quale è il digitale, in continuo divenire che impatta direttamente sulla società. Occorre dunque intervenire in tempo per delineare un confine, sul piano pratico, etico e giuridico, entro il quale circoscrivere l'implementazione e l'uso delle tecnologie. In caso contrario, gli effetti sarebbero allarmanti, senza possibilità di tutelare individui e società, restando impotenti di fronte alle tecnologie.

La Comunità Europea, a partire dal 2021²⁷⁶, sta cercando di indirizzare i Paesi membri attraverso la definizione di una politica comune volta a rendere l'AI «*trustworthy*», intesa ovvero come l'abilità di rispettare le aspettative di utenti e portatori di interesse in maniera affidabile, sicura ed etica²⁷⁷. L'esigenza avvertita è dunque quella di una maggiore regolarizzazione e standardizzazione delle caratteristiche e dell'uso dell'AI che trova, iniziale, risposta nel Regolamento sull'Intelligenza

²⁷² Conosciuto anche come Forum di Davos, ha avuto luogo nell'omonima città svizzera dal 15 al 19 Gennaio 2024.

²⁷³ Marc Benioff, CEO di Salesforce al Forum di Davos. Cfr. Giuliana Ferraino, «Il Forum di Davos, in diretta», *Corriere della Sera*, 18 Gennaio 2024, https://www.corriere.it/economia/finanza/diretta-live/24_gennaio_18/forum-davos-diretta-herzog-prima-una-soluzione-politica-vogliamo-certezza-pace-d0ee8932-b5cf-11ee-8131-eee5ee6b3690.shtml?refresh_ce.

²⁷⁴ Sara Della Piazza, «Il caso Cambridge Analytica», *DirittoConsenso*, 21 Dicembre 2021, <https://www.dirittoconsenso.it/2021/12/21/il-caso-cambridge-analytica/>.

²⁷⁵ Federico Rubini, «Il Forum di Davos, in diretta», *Corriere della Sera*, 18 Gennaio 2024, https://www.corriere.it/economia/finanza/diretta-live/24_gennaio_18/forum-davos-diretta-herzog-prima-una-soluzione-politica-vogliamo-certezza-pace-d0ee8932-b5cf-11ee-8131-eee5ee6b3690.shtml?refresh_ce.

²⁷⁶ Patrick Bezombes, «JTC 21 Preliminary views on ai act timeline and horizontal/vertical integration», <https://www.eurocontrol.int/sites/default/files/2023-05/20230420-flyai-forum-session3-cen-celecec-jtc-21-preliminary-views.pdf>

²⁷⁷ Commissione Europea, *Artificial Intelligence Act* (2023), p. 2, [https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI\(2021\)698792_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf)

Artificiale (*Artificial Intelligence Act* o AI Act), proposto dalla Commissione europea nel 2021²⁷⁸. Oggetto dell'AI Act è la minimizzazione del rischio (*risk-based approach*) connesso con l'impiego dell'intelligenza artificiale. Il documento individua quattro diverse classi di rischio:

1. Sistemi dal rischio inaccettabile;
2. Sistemi ad alto rischio;
3. Sistemi a rischio limitato;
4. Sistemi con rischio basso o minimo.

1. **Sistemi dal rischio inaccettabile.** Si tratta di sistemi di AI vietati il cui uso viola i diritti fondamentali. Sono inclusi in questa categoria sistemi che presentano «*un elevato potenziale in termini di manipolazione delle persone tramite tecniche subliminali, [...] oppure di sfruttamento delle vulnerabilità di specifici gruppi [...]*» e il ricorso (con alcune eccezioni) a sistemi di identificazione biometrica in “tempo reale” in luoghi accessibili al pubblico.
2. **Sistemi ad alto rischio.** Rientrano in questa categoria sistemi che costituiscono un elevato rischio per la salute e la sicurezza o per i diritti fondamentali degli individui che possono essere immessi sul mercato europeo previo il rispetto a determinati requisiti obbligatori. Nella categoria viene elaborata una diversificazione tra: sistemi di AI impiegati come componenti di sicurezza e altri sistemi di AI che presentano implicazioni con la sfera dei diritti fondamentali²⁷⁹.
3. **Sistemi a rischio limitato.** Si tratta di sistemi che interagiscono con gli umani, come ad esempio chatbot e assistenti virtuali, in grado di generare o manipolare contenuti multimediali e che sono soggetti al rispetto di obblighi di trasparenza.
4. **Sistemi con rischio basso o minimo.** Quest'ultima categoria comprende tutti i sistemi di AI il cui impiego presenta un rischio minimo per gli individui. Per tale ragione, possono essere sviluppati e impiegati in UE senza particolari obbligazioni legali seppure si prevede la creazione di codici di condotta a cui far aderire, su base volontaria, i provider di tali sistemi²⁸⁰.

Gli organismi di standardizzazione dell'UE, CEN (Comitato europeo di normalizzazione), CENELEC (Comitato europeo di normalizzazione elettrotecnica) e Joint Technical Committee

²⁷⁸ Commissione Europea, COM/2021/206 *Proposta di Regolamento del parlamento europeo e del consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (legge sull'intelligenza artificiale) e modifica alcuni atti legislativi dell'unione* (2021), https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0006.02/DOC_1&format=PDF.

²⁷⁹ Cfr. Commissione Europea, *Allegato III Sistemi di IA ad alto rischio di cui all'articolo 6, paragrafo 2* in COM/2021/206 *Proposta di Regolamento del parlamento europeo e del consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (legge sull'intelligenza artificiale) e modifica alcuni atti legislativi dell'unione* (2021), https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0006.02/DOC_2&format=PDF.

²⁸⁰ A differenza dei sistemi ad alto rischio per i quali è un obbligo.

Special Advisory Group sono stati incaricati di avviare il processo di analisi²⁸¹ di standard internazionali esistenti al fine di verificarne l'attuabilità nel quadro dell'Unione. Ciascun standard è stato analizzato in virtù di un possibile allineamento all'AI Act (*Alignment with the AI Act*) la cui prerogativa è quella di assicurare che i sistemi di AI adottati in Unione Europea rispettino le leggi e i valori²⁸² che la guidano e la caratterizzano. Per questa attività i sistemi di AI sono stati esaminati²⁸³ sulla base di dieci criteri individuati all'interno dell'Allegato 1 della Richiesta di Standardizzazione a supporto dell'AI Act²⁸⁴ e così distinti:

- I. **Risk management.** Il *risk management* è finalizzato a misurare e valutare il rischio connesso con i sistemi di AI prendendo in esame tre categorie principali: *risk assessment*, *testing* e *risk management*. Sono criteri intrinsecamente connessi tra loro per fornire una valutazione complessiva circa la fattibilità e sostenibilità del rischio legato all'AI.
- II. **Data quality and governance.** Rispettivamente, per qualità di dati si fa riferimento agli attributi che risultano essere rilevanti «*to the risk-oriented nature of the regulation, notably their relevance, representativeness, correctness and completeness, as well as their statistical properties*»²⁸⁵ in considerazione del contesto in cui vengono sviluppati i sistemi di intelligenza artificiale. La *data governance* fa riferimento invece alle procedure che investono le fasi di acquisizione, collezione, gestione e impiego dei dati durante il ciclo di vita dei sistemi con particolare attenzione alle modalità con cui avvengono tali processi.
- III. **Technical documentation, record keeping.** Mantenere traccia delle operazioni legate con i sistemi di AI, in particolare di quelli con rischio elevato, è fondamentale per avere consapevolezza di ciò che è stato compiuto durante l'intero ciclo di vita dei sistemi. Specificatamente per i sistemi a rischio elevato, la registrazione deve comprendere: (1) periodo, in data e ora, di utilizzo del sistema; (2) banca dati di riferimento per la verifica

²⁸¹ Soler Garrido et al., EUR 31518 EN *Analysis of the preliminary AI standardization work plan in support of the AI Act* (Publications Office of the European Union, Luxembourg, 2023), ISSN 1831-9424, <https://op.europa.eu/en/publication-detail/-/publication/b14d9c86-faa3-11ed-a05c-01aa75ed71a1/language-en>.

²⁸² Alessio Tartaro, "Towards European Standards supporting the AI Act: alignment challenges on the path to Trustworthy AI," *Proceedings of the AISB Convention 2023*, (2023), <https://aisb.org.uk/wp-content/uploads/2023/05/aisb2023.pdf>

²⁸³ La Commissione Europea ha difatti incaricato l'European Standardisation Organisations (ESO), a stilare una lista di standard necessari a soddisfare i requisiti per una intelligenza artificiale sicura e affidabile. Cfr. Soler Garrido et al., EUR 31518 EN *Analysis of the preliminary AI standardization work plan in support of the AI Act*, *cit.*, <https://op.europa.eu/en/publication-detail/-/publication/b14d9c86-faa3-11ed-a05c-01aa75ed71a1/language-en>.

²⁸⁴ Commissione Europea, *Draft standardisation request to the European Standardisation Organisations in support of safe and trustworthy artificial intelligence* (2022), <https://ec.europa.eu/docsroom/documents/52376>.

²⁸⁵ Soler Garrido et al., EUR 31518 EN *Analysis of the preliminary AI standardization work plan in support of the AI Act*, *cit.*, p. 8. <https://op.europa.eu/en/publication-detail/-/publication/b14d9c86-faa3-11ed-a05c-01aa75ed71a1/language-en>.

dei dati in input; (3) dati in input; (4) identificativo delle persone che partecipano alla verifica dei risultati²⁸⁶.

- IV. **Transparency and provision of information to users.** I sistemi dovrebbero essere sempre trasparenti nei confronti degli utenti e informarli dei rischi e delle modalità attraverso cui vengono collezionati i propri dati. La comunicazione deve essere quanto più chiara e semplice possibile affinché possa essere facilmente compresa ad ogni livello di impiego.
- V. **Human oversight.** Tale caratteristica si riferisce alla necessità di sviluppare i sistemi di AI consentendo la supervisione umana per l'intera durata dell'utilizzo. La supervisione umana deve essere sempre ammissibile per minimizzare il rischio connesso con l'acquisizione, da parte dei sistemi, di bias automatici che possono costituire un potenziale pericolo per il contesto di applicazione.
- VI. **Accuracy.** L'accuratezza misura la capacità dei sistemi di AI a eseguire efficacemente le attività per cui sono stati implementati. La valutazione legata al livello di accuratezza deve essere rintracciabile nelle istruzioni d'accompagnamento ai sistemi.
- VII. **Robustness.** Connessa con l'accuratezza, la caratteristica della robustezza indica il grado di resilienza dei sistemi verso errori, bug o altre eventualità che possono inficiarne il funzionamento.
- VIII. **Cybersecurity.** In merito alla cibersecurity viene preso in esame la capacità dei sistemi a resistere ad attacchi o tentativi volti ad alterarne l'uso. In tal senso vengono sono comprese ogni genere di soluzioni, tecniche e organizzative, finalizzate a minimizzare i rischi connessi alla violazione dei sistemi.
- IX. **Quality management.** Lo sviluppo di un sistema di gestione della qualità si rende necessario, in particolare per i fornitori di AI a rischio elevato, a garantire la conformità al Regolamento. Tali sistemi sono finalizzati a valutare l'impatto dell'intelligenza artificiale all'interno dell'organizzazione tenendo conto, nel loro insieme, delle caratteristiche sopra descritte. A tale scopo, il sistema è documentato delle procedure, politiche nonché istruzioni implicate nello sviluppo e adozione dei sistemi di AI.
- X. **Conformity assessment.** La valutazione di conformità è un processo obbligatorio a cui i fornitori di sistemi a rischio elevato devono attenersi prima di immettere sul mercato il prodotto. L'articolo 43 comma 1 del Regolamento disciplina le modalità attraverso cui la certificazione può essere ottenuta:

²⁸⁶ Commissione Europea, COM/2021/206 *Proposta di Regolamento del parlamento europeo e del consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (legge sull'intelligenza artificiale) e modifica alcuni atti legislativi dell'unione, cit., art. 53, https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0006.02/DOC_1&format=PDF.*

- i. tramite controllo interno come delineato nell'allegato VI al Regolamento;
- ii. tramite valutazione del sistema di gestione della qualità e della documentazione tecnica da parte di un organismo certificato secondo i criteri dell'allegato VII al Regolamento²⁸⁷.

Con l'obiettivo di delineare il quadro regolamentare attuale e il possibile scenario applicativo, la Commissione Tecnica ha effettuato un attento confronto tra i requisiti imposti da ciascun criterio e gli standard disponibili. Il processo è particolarmente delicato poiché è finalizzato all'adozione, nel territorio comunitario, di standard definiti a livello internazionale e, come tale, richiede un'attenta valutazione degli interessi posti in gioco. Nel piano di lavoro, la Commissione Tecnica ha individuato e analizzato, per ciascuna delle caratteristiche sopra descritte, gli standard attualmente in vigore definendone gli aspetti attuabili e le criticità. Il report presenta un quadro molto variegato in cui talvolta, seppure con qualche forzatura, un allineamento tra diversi standard sembra realizzabile. È questo il caso, ad esempio, dei requisiti di robustezza e accuratezza (art. 15) per i quali sono presi a riferimento standard di alto livello come ISO/IEC 22989 sui concetti e la terminologia o ISO/IEC 24028 sulle caratteristiche di affidabilità. Con un taglio specifico per la robustezza, invece, è stata richiamata la serie ISO/IEC 24029, funzionale allo sviluppo di modelli valutativi della robustezza attraverso una introduttiva panoramica – parte I – sulle metriche e misure attualmente disponibili per la valutazione delle reti neurali e – parte II – definendo le specifiche tecniche per lo sviluppo di metodi di verifica. Lo standard ISO/IEC AWI TS 29119-11, in aggiunta, definisce delle linee guida per testare efficacemente i sistemi basati sull'AI mentre la ISO/IEC AWI TS 17847 si focalizza maggiormente sugli aspetti inerenti alla validazione e verifica degli stessi. Tuttavia, l'individuazione di standard coerenti con i requisiti dell'AI Act, non impedisce la definizione in futuro di standard europei specialmente laddove, come nell'esempio riportato, l'attenzione è rivolta verso un quadro tecnico più ampio e generale. Il lavoro svolto dalla CEN-CENELEC JTC 21 rappresenta un momento chiave nel processo di standardizzazione dell'intelligenza artificiale poiché indica un'apertura dell'Europa verso l'internazionalizzazione. La determinazione ad attuare un confronto con altri contesti produce un effetto positivo sull'intero ecosistema digitale, ponendo le basi per una collaborazione e condivisione di conoscenza e competenze. È questo, d'altronde, l'obiettivo della Dichiarazione di Bletchley²⁸⁸, siglata nel novembre del 2023 da 28 paesi e comprendendo l'Unione Europea, Regno Unito, Cina e

²⁸⁷ Commissione Europea, COM/2021/206 *Proposta di Regolamento del parlamento europeo e del consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (legge sull'intelligenza artificiale) e modifica alcuni atti legislativi dell'unione, cit.*, art. 43, comma 1, https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0006.02/DOC_1&format=PDF.

²⁸⁸ Giulia Amadeo, "Dichiarazione di Bletchley: il via alla cooperazione internazionale per un'intelligenza artificiale sicura," *Altalex*, 19 Novembre 2023, <https://www.altalex.com/documents/news/2023/11/19/dichiarazione-di-bletchley-via-a-cooperazione-internazionale-per-intelligenza-artificiale-sicura>.

Stati Uniti. La Dichiarazione stabilisce un accordo tra le principali potenze in materia di intelligenza artificiale in merito ai rischi implicati con lo sviluppo e l'uso di tali sistemi e, dunque, ammettendo la necessità di lavorare insieme per indirizzarla a beneficio della comunità internazionale²⁸⁹.

Le aree di intervento del Regolamento sono volte, dunque, a definire criteri standard per lo sviluppo di sistemi basati sull'intelligenza artificiale nel rispetto dei diritti fondamentali dell'uomo così come espressi nella Carta dei Diritti Fondamentali dell'Unione Europea²⁹⁰. All'interno del documento, dove i diritti sono classificati nelle sei categorie di dignità, libertà, uguaglianza, solidarietà, diritti civili e giustizia, l'aspetto culturale è richiamato marginalmente, più che altro come rispetto alle diversità (art. 22²⁹¹) e al diritto a vivere una vita dignitosa partecipando socialmente e culturalmente (art. 25²⁹²). Non è dunque presente un esplicito richiamo sull'impiego di tali tecnologie al patrimonio culturale che, per sua natura, connota gli aspetti educativi e sociologici umani. Su quest'ultimo aspetto si focalizza invece la Raccomandazione sull'Etica dell'Intelligenza Artificiale, adottata dall'UNESCO²⁹³ nel novembre del 2021. La Raccomandazione tratta le implicazioni etiche coinvolte con i sistemi di AI e l'applicazione sui processi decisionali nei diversi settori (culturali, sociali, economici etc.). L'obiettivo del presente documento è quello di indirizzare le tecnologie di AI che hanno una maggiore rilevanza etica connesse con la capacità di processare dati e informazioni che possono impattare direttamente sulla sicurezza, la salute, l'ambiente, la società etc. Particolare attenzione è posta verso i sistemi basati sul Machine e Deep Learning all'interno dei settori di interesse per l'UNESCO e oggetto di un preliminare studio, nel 2019, da parte della UNESCO World Commission on Ethics of Scientific Knowledge and Technology (COMEST). L'ambito culturale è qui inteso nella sua accezione *tout court* in relazione con le tecnologie di intelligenza artificiale, laddove quest'ultime «*can also lead to an increase concentration of supply of cultural content, data, markets and income in the hands of only a few actors, with potential negative implications for the diversity and pluralism of languages, media, cultural expressions, participation and equality*»²⁹⁴.

L'accento è qui posto, in particolare, sulle conseguenze sul piano sociologico manifestando l'interesse vero gli effetti su individui e società di un uso improprio dell'AI sul patrimonio culturale. È questo un aspetto di vitale importanza all'interno della società, digitale quanto analogica, in cui la privazione

²⁸⁹ Luisa Franchina, "Dichiarazione di Bletchley, svolta per lo sviluppo responsabile dell'intelligenza artificiale: ecco perché", *Cybersecurity 360*, 8 Novembre 2023, <https://www.cybersecurity360.it/cultura-cyber/dichiarazione-di-bletchley-svolta-per-lo-sviluppo-responsabile-dellintelligenza-artificiale-ecco-perche/>.

²⁹⁰ Commissione Europea, 2000/C 364/01, *Charter of Fundamental Rights of the European Union* (2020), https://www.europarl.europa.eu/charter/pdf/text_en.pdf.

²⁹¹ «*Cultural, Religious and linguistic diversity – shall be respected*».

²⁹² «*The rights of the elderly – to live a life of dignity and to participate in social and cultural life*».

²⁹³ United Nations for Education, Science and Culture Organization (UNESCO), *Recommendation on the Ethics of Artificial Intelligence* (2021), <https://unesdoc.unesco.org/ark:/48223/pf0000380455>.

²⁹⁴ United Nations for Education, Science and Culture Organization (UNESCO), *Recommendation on the Ethics of Artificial Intelligence, cit.*, p.11, <https://unesdoc.unesco.org/ark:/48223/pf0000380455>.

o manipolazione del patrimonio culturale riflette concetti di predominio, razziale e sociale. Il riconoscimento dei pericoli connessi con l'applicazione di AI sul patrimonio culturale costituisce uno dei pilastri delle Raccomandazioni UNESCO, un framework²⁹⁵ di valori, principi e buone pratiche messo a disposizione degli Stati nella definizione di politiche interne inerenti l'AI conformi con la legislazione internazionale. Fra i diversi contesti presi ad esame, la cultura è oggetto della *policy area* 7²⁹⁶ in cui si incoraggia gli Stati membri ad adottare (artt. 94-95) appropriati strumenti basati sull'AI per la gestione, valorizzazione, e conservazione del patrimonio culturale, materiale e immateriale, e salvaguardia delle comunità culturali a rischio di estinzione. Gli artt. 97 e 98 sono invece indirizzati a ravvisare i pericoli connessi con l'offuscamento e la privazione del patrimonio culturale e dunque sulla necessità di promuovere attività volte a migliorarne la visibilità e la scoperta. È importante che anche gli archivi, musei, gallerie e biblioteche (art. 100), in qualità di depositari della conoscenza, siano sostenuti dai propri Governi a sfruttare le tecnologie per valorizzare il patrimonio culturale di appartenenza e migliorare la qualità dei propri servizi cosicché da ottimizzare l'accesso alle risorse da parte degli utenti.

Un altro degli aspetti connessi con la standardizzazione dell'intelligenza da tenere in considerazione artificiale, è il processo evolutivo delle tecnologie che per la rapidità del suo corso implica un grado elevato di dinamismo a livello regolamentatore. Difatti, se da una parte la standardizzazione agevola la creazione e la proliferazione di appropriate risorse, dall'altra invece comporta il rischio di promuovere l'utilizzo di tecnologie datate e obsolete²⁹⁷. Se da una parte, limitare il progresso tecnologico sembra di fatti un'utopia, occorre tuttavia definire un contesto etico e giuridico entro il quale esso debba avvenire senza nuocere alla sfera dell'individuo, anche e soprattutto in rapporto col territorio.

Attraverso le Raccomandazioni, l'UNESCO ha contribuito ad avviare un processo volto a far emergere la consapevolezza in merito all'impatto che l'uso delle tecnologie può avere sul patrimonio culturale mondiale. Maggiori difficoltà sono presenti invece sul piano giuridico dove le finalità dei diversi governi sembrano non collimare tra loro. Persino nell'Unione Europea, l'approvazione dello stesso AI Act sembra tutt'altro che scontata con la contrapposizione di Italia, Francia e Germania per

²⁹⁵ United Nations for Education, Science and Culture Organization (UNESCO), *Recommendation on the Ethics of Artificial Intelligence*, cit., pp.14-15, <https://unesdoc.unesco.org/ark:/48223/pf0000380455>.

²⁹⁶ United Nations for Education, Science and Culture Organization (UNESCO), *Recommendation on the Ethics of Artificial Intelligence*, cit., artt. 94-100, <https://unesdoc.unesco.org/ark:/48223/pf0000380455>.

²⁹⁷ Swann, GM Peter, "The Economics of Standardization: An Update," *Report for the UK Department of Business, Innovation and Skills (BIS)*, 2010, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/32444/10-1135-economics-of-standardization-update.pdf.

deregolamentare i nuovi modelli di AI generativa²⁹⁸. L'accordo trilaterale prevede un modello di auto-regolamentazione dei sistemi di intelligenza artificiale promuovendo, in capo agli sviluppatori di AI, una responsabilità settoriale. Si tratta dunque di una concezione più elastica rispetto a quanto proposto con l'AI Act che promuove, invece, un approccio basato sul rischio di tali sistemi.

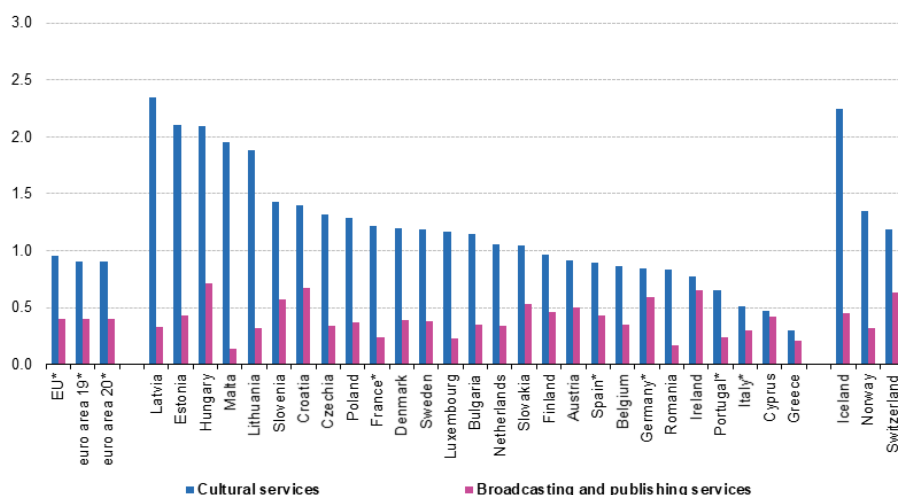
Nella sfera culturale, la precarietà dello scenario regolamentatore si riversa inevitabilmente sui contenuti che finiscono per essere oggetto di sfruttamento e privatizzazione. A partire già dal mercato tecnologico, si attesta una forte dipendenza dell'UE nei confronti di aziende provenienti da paesi esterni per la gestione dei propri dati²⁹⁹. La dipendenza da fornitori esterni all'eurozona per la gestione del proprio patrimonio culturale rappresenta una sconfitta per l'intera Unione Europea che annovera, tra i propri Paesi, il maggior numero di siti censiti come Patrimonio dell'Umanità dall'UNESCO³⁰⁰. La situazione diviene ancora più critica nel contesto nazionale dove l'Italia, nonostante detenga il maggior numero di siti censiti dall'UNESCO – ben 59 tra beni naturali, culturali e patrimonio immateriale –, risulta agli ultimi posti in Europa per fondi investiti nella cultura come rilevato nell'indagine Eurostat in merito alla percentuale della spesa dei Governi europei allocata nei servizi per la cultura, radiotelevisivi e l'editoria (Figura 32).

²⁹⁸ Francesco Niola, "L'Italia vuole deregolamentare l'IA generativa, ecco come," *Agenda Digitale*, 27 Novembre 2023, <https://www.agendadigitale.eu/sicurezza/privacy/litalia-vuole-deregolamentare-lia-generativa-ecco-come/>.

²⁹⁹ Parlamento Europeo, "Artificial intelligence in the context of cultural heritage and museums: Complex challenges and new opportunities," *cit.*, p. 6, [https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/747120/EPRS_BRI\(2023\)747120_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/747120/EPRS_BRI(2023)747120_EN.pdf)

³⁰⁰ United Nations for Education, Science and Culture Organization (UNESCO), "World Heritage Online Map Platform," <https://experience.arcgis.com/experience/aeae427b3c5045a6a4c72a93ab1d1280>

General government expenditure on cultural services, broadcasting and publishing services, 2021 (% of total expenditure)



* provisional
Source: Eurostat (gov_10a_exp)

eurostat

Figura 32: Dati risultati dall'indagine Eurostat in merito alla spesa dei Governi europei per i servizi culturali, radiotelevisivi e l'editoria³⁰¹.

I risultati emersi dall'indagine delineano una situazione infelice per l'Italia, in cui soltanto il 3% della spesa pubblica è stato investito per la cultura nel biennio 2020-2021. Il dato può essere ulteriormente letto alla stregua degli effetti della pandemia da COVID-19 che hanno evidenziato la necessità di sfruttare le tecnologie e il digitale per promuovere efficacemente il patrimonio culturale. Nella situazione di *lockdown* con la limitazione degli spostamenti e di attività in presenza, l'ecosistema digitale ha rappresentato l'unico potenziale strumento in grado di far fruire del patrimonio culturale e, come tale, un asset strategico per imprese e governi. Questi ultimi, in particolare, hanno la responsabilità di far sì che il patrimonio culturale non venga sfruttato come mero tornaconto economico, bensì che possa costituire un vantaggioso per le comunità, per la crescita dei territori, come aggregatore sociale e per la promozione di cultura. Si tratta di un compito di vitale importanza espressamente richiamato, per il contesto italiano, all'articolo 1, comma 3 del Codice dei beni culturali e del paesaggio in cui si legge che «Lo Stato, le regioni, le città metropolitane, le province e i comuni assicurano e sostengono la conservazione del patrimonio culturale e ne favoriscono la pubblica fruizione e valorizzazione»³⁰² ricordando che la valorizzazione, come indicato al comma 2

³⁰¹ Eurostat, "Government expenditure on cultural, broadcasting and publishing services," Ultimo accesso 7 Gennaio 2024, https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Government_expenditure_on_cultural_broadcasting_and_publishing_services#General_government_expenditure_on_cultural_services.2C_broadcasting_and_publishing_services.

³⁰² D.Lgs 42/2004 *Codice dei beni culturali e del paesaggio*, cit., art. 1, comma 3, <https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:decreto.legislativo:2004-01-22;42!vig=>.

del medesimo articolo, concorre insieme alla tutela «a preservare la memoria della comunità nazionale e del suo territorio e a promuovere lo sviluppo della cultura»³⁰³.

L'approccio umano verso l'intelligenza artificiale

Il tema sulla standardizzazione dell'intelligenza ha evidenziato la necessità di investire sulla formazione digitale nei diversi settori per far fronte ai pericoli derivanti da un uso inconsapevole delle tecnologie. Le società sono chiamate a rispondere alla sfida promuovendo la formazione scolastica e universitaria unitamente a percorsi di *upskilling* per rafforzare le competenze digitali e favorire un'educazione orientata ad un uso consapevole dell'intelligenza artificiale. Ciò si rende necessario per contrastare l'idea, nociva, che ponga l'intelligenza artificiale come risolutore assoluto di ogni problema. Gli stessi modelli di AI generativi ad oggi sviluppati e largamente impiegati, come ChatGPT, costituiscono un valido strumento solamente laddove vi sia un utilizzo coscienzioso e consapevole senza pretendere che esso possa sostituirsi all'attività umana. Sono noti, a tal proposito, casi in cui l'uso inconsapevole di AI generativa ha portato a conclusioni negative. A tal proposito si può citare l'episodio avvenuto a New York in una causa contro la compagnia aerea Avianca dove, a seguito di precedenti citati dall'accusa che non hanno avuto alcun riscontro, l'avvocato ha confessato di aver utilizzato, inadeguatamente, ChatGPT per rintracciare dei casi senza verificarne l'autenticità³⁰⁴. Circostanze simili avvengono laddove tali sistemi sono impostati per decidere autonomamente senza una "supervisione" umana in grado di contestualizzarne l'applicazione. Occorre, difatti, porre l'attenzione su questo aspetto e demistificare l'idea che l'intelligenza artificiale possa risolvere, autonomamente, i problemi dell'uomo. Essa costituisce un valido aiuto per chi, conscio delle opportunità e dei rischi legati al dominio di applicazione, è in grado di ragionare sui risultati ottenuti e di contestualizzarli andando oltre ai singoli dati ottenuti. Ciò si riflette in particolare su quei processi che richiedono un alto livello di astrazione che Kahneman³⁰⁵ definisce "pensieri lenti". Tale definizione sottintende alle attività per le quali è necessaria una maggiore capacità di ragionamento contrapposte ai "pensieri veloci" che richiedono, invece, una maggiore reattività. Nell'elaborazione delle immagini, è questo il caso del task di *Visual Commonsense Reasoning* (VCR)³⁰⁶ in cui viene sottoposto alla macchina il compito di comprendere, oltre che di riconoscere, ciò che è contenuto nelle immagini. È un'attività che richiede un certo tipo di ragionamento e, a differenza di task come la classificazione o il riconoscimento, non può fornire un immediato riscontro. Prendendo ad esame

³⁰³ D.Lgs 42/2004 *Codice dei beni culturali e del paesaggio*, cit., art. 1, comma 2, <https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:decreto.legislativo:2004-01-22;42!vig=>.

³⁰⁴ Benjamin Weiser, "Here's What Happens When Your Lawyer Uses ChatGPT," *New York Times*, 27 Maggio 2023, <https://www.nytimes.com/2023/05/27/nyregion/avianca-airline-lawsuit-chatgpt.html>.

³⁰⁵ Daniel Kahneman, *Thinking Fast and Slow* (United Kingdom: Penguin Books 2012). ISBN: 978-0141033570.

³⁰⁶ Rowan Zellers et al., "From Recognition to Cognition: Visual Commonsense Reasoning," (2019). arXiv:1811.10830.

la Figura 33, l'intelligenza artificiale deve essere in grado di comprendere e dare una giustificazione razionale delle dinamiche che entrano in gioco dalle situazioni catturate dall'immagine. Nel caso specifico del ristorante, l'AI, oltre a dare la risposta esatta sull'azione in corso d'opera, deve essere abile a giustificare la scelta della risposta, ovvero a ragionare sugli elementi che hanno portato a prendere tale decisione.

Why is [person4] pointing at [person1]?

- a) He is telling [person3] that [person1] ordered the pancakes.
- b) He just told a joke.
- c) He is feeling accusatory towards [person1].
- d) He is giving [person1] directions.

I chose a) because...

- a) [person1] has the pancakes in front of him.
- b) [person4] is taking everyone's order and asked for clarification.
- c) [person3] is looking at the pancakes and both she and [person2] are smiling slightly.
- d) [person3] is delivering food to the table, and she might not know whose order is whose.

How did [person2] get the money that's in front of her?

- a) [person2] is selling things on the street.
- b) [person2] earned this money playing music.
- c) She may work jobs for the mafia.
- d) She won money playing poker.

I chose b) because...

- a) She is playing guitar for money.
- b) [person2] is a professional musician in an orchestra.
- c) [person2] and [person1] are both holding instruments, and were probably busking for that money.
- d) [person1] is putting money in [person2]'s tip jar, while she plays music.

Figura 33: Task di Visual Commonsense Reasoning³⁰⁷.

La complessità del task è determinata dalla capacità di pensiero richiesto per il quale, essendo necessario un alto livello di astrazione e ragionamento, l'uomo dimostra maggiore precisione rispetto alle macchine. Il grafico nella Figura 34 conferma l'attuale situazione di vantaggio della capacità di ragionamento dell'uomo (*Human Baseline*) rispetto all'AI sul task di VCR misurata tramite il Q→AR Score, la metrica che combina la capacità della macchina a selezionare la risposta esatta con l'abilità a motivarla razionalmente.

³⁰⁷ Rowan Zellers et al., "From Recognition to Cognition: Visual Commonsense Reasoning," *cit.*, p.1. arXiv:1811.10830.

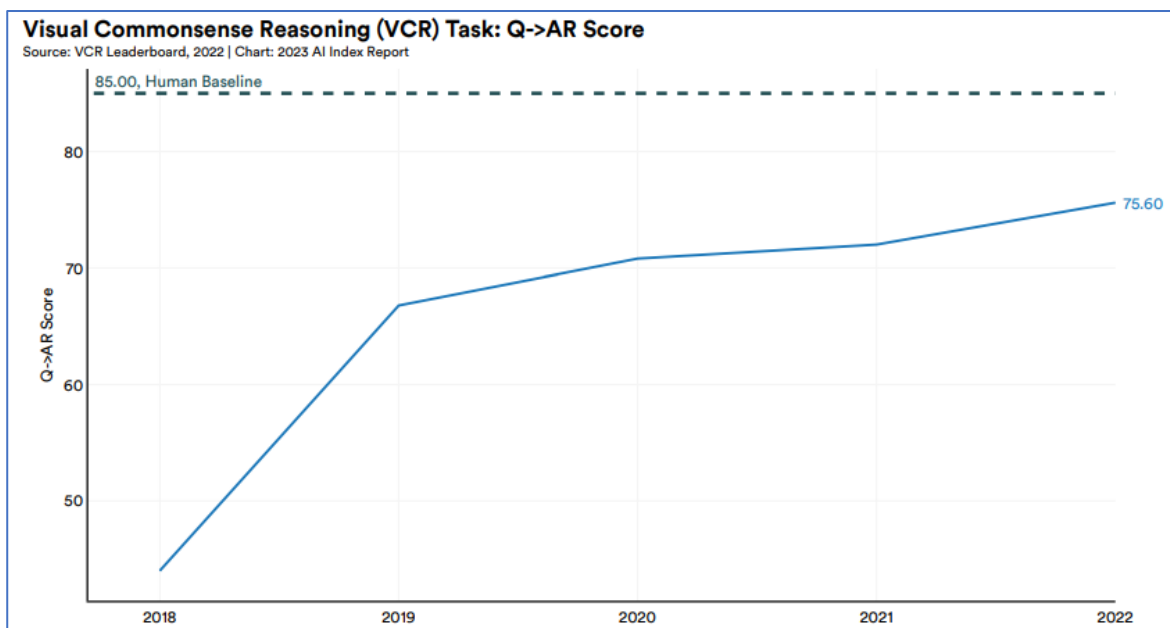


Figura 34: confronto del punteggio tra Baseline umana e intelligenza artificiale³⁰⁸.

Le attività come quelle riportate nell'esempio, devono far riflettere sulla demarcazione tra uomo e macchina e sulle potenzialità di quest'ultima come strumento per elaborare processi di rapida esecuzione e, così facendo, incrementare la produttività.

Oggi giorno si sta acquisendo maggiore consapevolezza a riguardo e, come testimoniato dall'AI Act, il trend di pensiero sull'intelligenza artificiale sta mutando: non occorre disporre di soli informatici in grado di sviluppare modelli, bensì integrare le conoscenze con esperti di dominio al fine di rendere le soluzioni realmente efficaci muovendole da un alto livello di generalità verso una maggiore settorialità. È quanto sta accadendo, ad esempio, nell'ambito medico dove vengono sviluppati un sempre più crescente numero di dispositivi, certificati³⁰⁹, basati sull'AI in grado di assistere i medici nella lettura delle immagini³¹⁰, nelle diagnosi sino ad arrivare alle operazioni chirurgiche³¹¹. Se in tale ambito la ricerca continua ad essere altamente proficua, altrettanto non può dirsi in merito al patrimonio culturale come testimonia il limitato numero di dataset di immagini disponibili in rete. Tale è, difatti, il risultato da un confronto effettuato³¹² sulla piattaforma Kaggle tra il numero di

³⁰⁸ Stanford University Human-Centered Artificial Intelligence, *Artificial Intelligence Index Report* (2023), p. 95. https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI_AI-Index-Report_2023.pdf

³⁰⁹ Allo stato attuale sono 161 i dispositivi approvati dalla *Food and Drug Administration*. Cfr. Food and Drug Administration (FDA), "Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices," ultima modifica 19 Ottobre 2023, <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices>. Ultimo accesso 23 Gennaio 2024.

³¹⁰ Pawan Kumar Mall et al., "A comprehensive review of deep neural networks for medical image processing: Recent developments and future opportunities," *Healthcare Analytics* 4, (2023). DOI: 10.1016/j.health.2023.100216.

³¹¹ Yiming Zhang, Ying Weng, and Jonathan Lund, "Applications of Explainable Artificial Intelligence in Diagnosis and Surgery," *Diagnostics* 12, n. 2 (2022). DOI: 10.3390/diagnostics12020237.

³¹² I dati sono aggiornati al 23 Gennaio 2024.

dataset di immagini disponibili in ambito sanitario – Figura 35 – (per i quali risultano disponibili oltre 3.700 dataset) e culturale – Figura 36 – (per i quali risultano disponibili solamente 52 dataset).

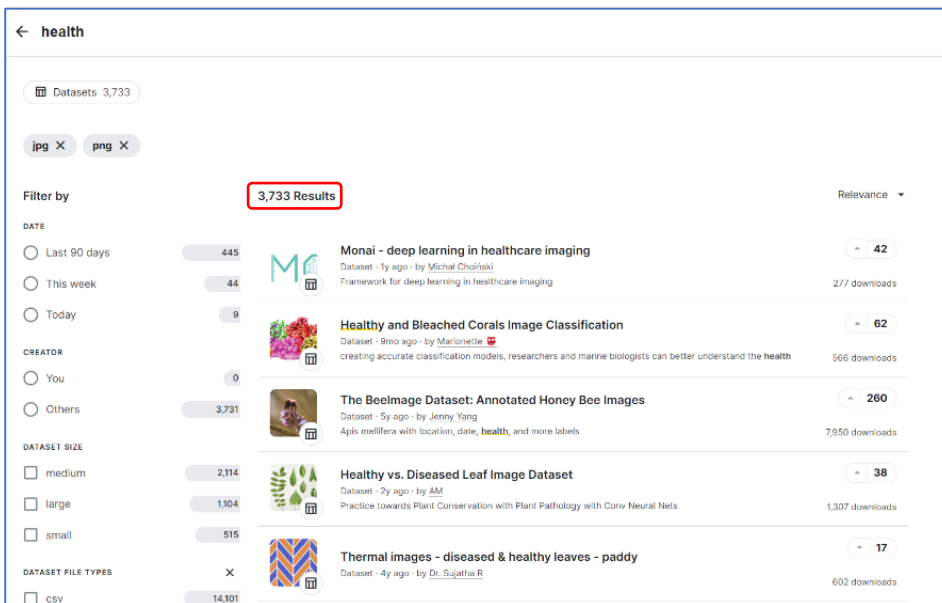


Figura 35: screenshot ricerca dataset immagini "health".

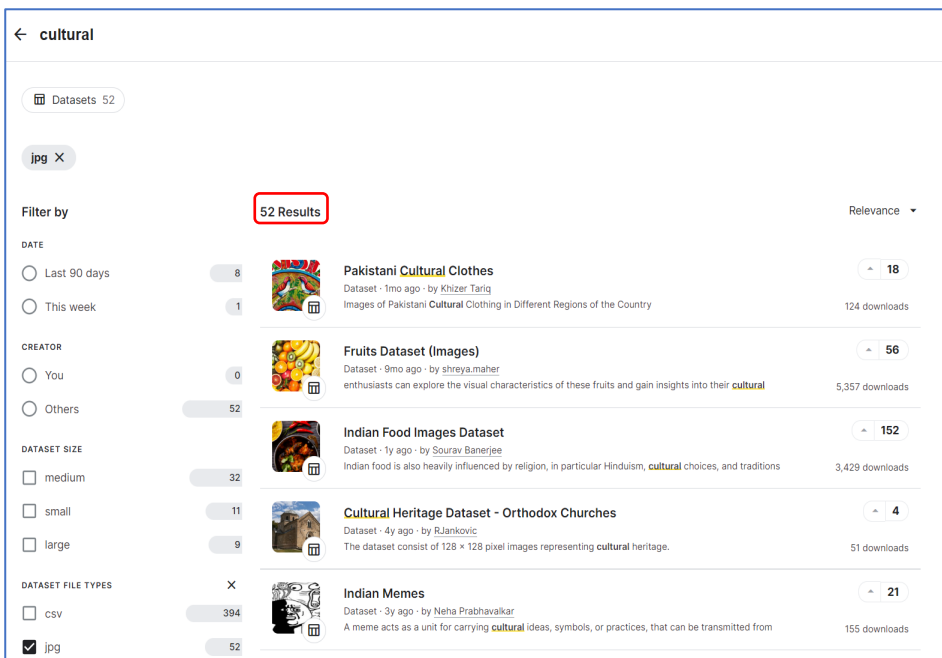


Figura 36: screenshot ricerca dataset immagini "cultural".

Il confronto mostra un risultato decisamente impari che sottolinea, ancora una volta, l'impegno e l'investimento profuso verso l'ambito sanitario rispetto a quello culturale. In merito a quest'ultimo,

in base all'attuale disponibilità in rete di dataset di immagini per l'applicazione di tecniche di intelligenza artificiale, è possibile avanzare ulteriori considerazioni:

1. i dataset attualmente esistenti mostrano un maggiore interesse verso il task di classificazione;
2. rispetto alla classificazione del patrimonio culturale, le categorie maggiormente prese ad esame sono i beni architettonici e il patrimonio immateriale per la componente enogastronomica e folkloristica;
3. specificatamente al patrimonio immateriale, si tratta di dataset contenenti immagini specifiche di culture medio-orientali, come ad esempio abiti tradizionali pakistani³¹³ o del Nepal³¹⁴, piatti tipici indiani³¹⁵;
4. in generale, esiste una netta predominanza di dataset aventi per oggetto la cultura medio-orientale rispetto a quella europea. Prendendo ad esame gli hub di HuggingFace e Kaggle e filtrando la ricerca per “*cultural*”, su un totale di 33 dataset realmente pertinenti al patrimonio culturale, 23 caratterizzano aspetti particolari della cultura medio orientale, 4 quella europea, soltanto uno l’America Latina e, infine, 5 hanno un tratto più generale prendendo a riferimento oggetti come dipinti e raffigurazioni non esclusive di determinate culture.

La carenza di dati che tale scenario presenta è sinonimo di uno scarso interesse verso il settore culturale in considerazione, oltretutto, dell'importanza che essi rivestono nei processi di addestramento e, per i quali, è indispensabile un accurato trattamento. Tale attività è finalizzata a certificare che i dati siano stati correttamente etichettati ed organizzati e, di conseguenza, è necessario l'intervento di esperti di dominio in grado di individuare eventuali problematicità e saperle correggere, anche autonomamente.

Tali sono dunque le *skill* alla base delle *Digital Humanities* (DH), una «*transdiscipline*»³¹⁶ in progressiva espansione nel panorama nazionale ed europeo con un numero sempre più elevato di corsi di studi attivati³¹⁷ che coniuga le metodologie informatiche con le discipline umanistiche. In

³¹³ Khizer Tariq, *Pakistani Cultural Clothes*, collezionato tramite Google Images e disponibile su Kaggle, Licenza MIT, <https://www.kaggle.com/datasets/khizertariq/pakistani-cultural-clothes>.

³¹⁴ Bimarsha Khanal, *Nepali Cultural Dress and Ornaments*, collezionato tramite Internet e disponibile su Kaggle, Licenza MIT, <https://www.kaggle.com/datasets/bimarshakhanal/nepali-cultural-dress-and-ornaments>.

³¹⁵ Sourav Banerjee, *Indian Food Images Dataset*, collezionato tramite Google Images e disponibile su Kaggle, <https://www.kaggle.com/datasets/iamsouravbanerjee/indian-food-images-dataset>.

³¹⁶ Pierre Mounier, “Manifeste des Digital Humanities,” *Journal des anthropologues*, n. 122-123 (2010), p. 450. DOI: 10.4000/jda.3652.

³¹⁷ Per il contesto italiano, il numero di percorsi di laurea magistrale nell'ambito delle *Digital Humanities* all'a.a. 2020/21 è pari al 14,3% (58 corsi su 406) contro il 10,6% registrato nell'a.a. 2010/2011. Cfr. Consorzio Interuniversitario AlmaLaurea, *XXIII Indagine Profilo dei Laureati 2020. Rapporto 2021* (2021), https://www.almalaurea.it/sites/default/files/2022-05/almalaurea_profilo_rapporto2021_0.pdf.

tale concezione, i profili in uscita dai percorsi di DH dovrebbero essere in grado, nei rispettivi settori disciplinari di appartenenza, di applicare al meglio le tecnologie a disposizione, ossia essere consapevoli delle modalità e degli obiettivi perseguibili intervenendo, laddove necessario, a correggere gli strumenti con la conoscenza di dominio. In un settore fortemente eterogeneo come quello culturale, ciò comporta necessariamente un'appropriata specializzazione di materia per far sì che le tecnologie siano sfruttate efficacemente per cogliere le singole sfaccettature che ciascun bene, culturale o naturale, presenta. Difatti, solamente conoscendo l'ambito di applicazione è possibile definire le corrette modalità attraverso cui impiegare gli strumenti e, di conseguenza, contestualizzare e dare valore ai risultati ottenuti. Processare un documento d'archivio e un manufatto richiedono competenze differenti che siano in grado di conoscerne le singole peculiarità e sapere come e quali tecnologie applicare. Ciò non deve escludere, tuttavia, la possibilità di mettere in relazione tra loro elementi che, seppur diversi sotto la forma, sono riconducibili alla medesima entità di patrimonio culturale. In tale circostanza è bene ribadire l'unità, sotto il profilo concettuale, del patrimonio culturale i cui diversi "settori", ognuno con le proprie peculiarità, sono intrinsecamente connessi tra loro.

Gli istituti culturali, quali depositari e promotori di cultura, sono tenuti a rispondere al bisogno collettivo di una maggiore fruibilità del patrimonio culturale, in particolare nell'ecosistema digitale in cui la possibilità di avere immediata disponibilità dei servizi accentua tale esigenza. In ambiente digitale, le immagini costituiscono uno dei principali formati documentali per la condivisione di materiale culturale. All'interno di queste, così come di dipinti, disegni etc., si cela una importante mole informativa che necessita di essere capitalizzata e condivisa per arricchire, soprattutto sotto il profilo descrittivo, e valorizzare il patrimonio culturale.

Partendo dall'analisi del dominio culturale e dello sviluppo tecnologico, il presente lavoro di tesi intende sottolineare i vantaggi per la promozione e disseminazione del patrimonio culturale conseguibili attraverso un'applicazione congiunta, e automatica, di strumenti e tecnologie differenti. A tale scopo, si è proceduto a delineare i principali punti di connessione tra le reti neurali convoluzionali e l'*International Image Interoperability Framework*, che permettono di estrarre e mettere in risalto elementi informativi, talvolta nascosti, presenti nelle immagini del patrimonio culturale. Il workflow ipotizzato ha l'obiettivo di mettere in evidenza la concreta opportunità di applicare, efficacemente, tecnologie quali le reti neurali convoluzionali al settore culturale per estrarre una conoscenza "latente" che possa essere interpretata, organizzata e rappresentata per valorizzare il patrimonio culturale. Per gli obiettivi posti in gioco, si tratta di un processo trasversale che implica necessariamente conoscenza del dominio e delle tecnologie da applicare per conseguire il risultato desiderato. Oltretutto, la formalizzazione degli elementi estraibili con standard di metadati o modelli

ontologici conformi consente agli istituti culturali di migliorare la qualità dei dati attraverso un “arricchimento semantico”³¹⁸ per permettere a studiosi e ricercatori, una maggiore cognizione delle risorse culturali. Ad aggiungere valore, sotto l’aspetto visivo, contribuiscono infine i sofisticati visualizzatori oggi disponibili che permettono ad archivi, musei e biblioteche, di dare un’ulteriore rilevanza alle proprie collezioni e di offrire agli utenti nuove modalità di accesso. Il risultato che ne consegue, è la possibilità di potenziare la scoperta del patrimonio culturale comportando, da un lato, un notevole miglioramento nella qualità dei servizi, dall’altro, un livellamento delle discrepanze e delle barriere culturali.

³¹⁸ Marcia Lei Zeng, “Semantic enrichment for enhancing LAM data and supporting digital humanities. Review article,” *El profesional de la información* 28, n. 1(2019). DOI : 10.3145/epi.2019.ene.03.

Bibliografia

- Acke, Lien, Kristel De Vis, Stijn Verwulgen, and Jouke Verlinden. "Survey and literature study to provide insights on the application of 3D technologies in objects conservation and restoration." *Journal of Cultural Heritage* 49, (2021). DOI: 10.1016/j.culher.2020.12.003.
- Addison, Howard, Eunbyung Park, and Wendy Kan. *ImageNet Object Localization Challenge* (2018). <https://www.kaggle.com/competitions/imagenet-object-localization-challenge/overview>.
- Agenzia per l'Italia Digitale (AGID). *Linee Guida sulla Formazione, Gestione e Conservazione dei Documenti Digitali* (2021), Allegato 3 "Certificazione di Processo." https://www.agid.gov.it/sites/default/files/repository_files/allegato_3_certificazione_di_processo.pdf.
- Alù, Angelo, and Alessandro Longo. "Cos'è il digital divide, nuova discriminazione sociale (e culturale)." *Agenda Digitale*, 13 Marzo 2020. <https://www.agendadigitale.eu/cultura-digitale/il-digital-divide-culturale-e-una-nuova-discriminazione-sociale/>.
- American Museum of Natural History. "Dublin Core to VRA Core Mapping." Ultimo accesso 23 Gennaio 2024. <https://sites.google.com/site/dctovrc/home>.
- Andreu, Javier, and Pablo Serrano. "Contributions of the digital photogrammetry and 3D modelling of Roman inscriptions to the reading of damaged tituli: An example from the Hispania Tarraconensis (Castiliscar, Saragossa)." *Digital Applications in Archaeology and Cultural Heritage* 12, (2021). DOI: 10.1016/j.daach.2019.e00091.
- Anna V. Guryanova, I. V. Smotrova, A. E. Makhovikov, and A. S. Koychubaev. "Socio-ethical Problems of the Digital Economy: Challenges and Risks." *Digital Transformation of the Economy: Challenges, Trends and New Opportunities. Advances in Intelligent Systems and Computing* 908, (2019). DOI: 10.1007/978-3-030-11367-4_9.
- Apache TIKA. "Apache TIKA" Ultimo accesso 23 Gennaio 2024. <https://tika.apache.org/1.20/formats.html>.
- Apache TIKA. "Formats." Ultimo accesso 23 Gennaio 2024. <https://tika.apache.org/1.20/formats.html>.
- Aproso, Alessio Palmero, and Giovanni Moretti. "Tint 2.0: an All-inclusive Suite for NLP in Italian." *Proceedings of the Fifth Italian Conference on Computational Linguistic CLiC-it*, 10. 2018. <https://books.openedition.org/aaccademia/3571?lang=it>.
- Archivio Centrale dello Stato. "E' online la digital library dell'ACS." Ultimo accesso 23 Gennaio 2024. <https://acs.cultura.gov.it/e-online-la-digital-library-dellacs/>.
- Archivio Centrale dello Stato. "Teca Digitale ACS." Ultimo accesso 23 Gennaio 2024. <https://tecadigitaleacs.cultura.gov.it/media/esplora>.
- Archivio Storico Università di Bologna. "Fascicolo di Pier Paolo Pasolini in IIF." Ultimo accesso 23 Gennaio 2024. https://archiviostorico.unibo.it/it/patrimonio-documentario/fascicolo-e-tesi-di-pier-paolo-pasolini/fascicolo-da-studente-di-pier-paolopasolini?view=ImageView&manifest=https%3A%2F%2Farchiviostorico.unibo.it%2Fiiif%2F%2Fmanifest%2Farchiviostorico%2Ffascicolo_e_tesi_di_pier_paolo_pasolini%2Ffascicolo_pasolini.json&canvas=https%3A%2F%2Farchiviostorico.unibo.it%2Fiiif%2F%2Fcanvas%2Farchiviostorico%40fascicolo_e_tesi_di_pier_paolo_pasolini%40fascicolo_pasolini%402b8ccc5f6a848ca713a2bcffd5900ae032620d40-01.jp2.json.

Archivio Storico Università di Bologna. "Patrimonio documentario." Ultimo accesso 23 Gennaio 2024. <https://archivistorico.unibo.it/it/patrimonio-documentario>.

Aseeva, Irina, and Vladimir Budanov, "Digitalization: potential risks for civil society," *Economic Annals – XXI* 186, n. 1 (2020). DOI: 10.21003/ea.V186-05.

AssemblyAI. "Documentation." Ultimo accesso 23 Gennaio 2024. <https://www.assemblyai.com/docs/>.

Associazione Italiana Biblioteche. "Osservazioni sul DM 11 aprile 2023, n. 161." Ultimo accesso 23 Gennaio 2024. <https://www.aib.it/notizie/osservazioni-sul-dm-11-aprile-2023-n-161/>.

Aurora. "Aurora Project." Ultimo accesso 17 Gennaio 2024. <https://www.aurora-euproject.eu/>.

Aussenac-Gilles, Nathalie, Brigitte Biebow, and Sylvie Szulman. "Revisiting ontology design: A methodology based on corpus analysis." *Knowledge Engineering and Knowledge Management Methods, Models, and Tools. (EKAW 2000) 1937*, (2000). DOI: 10.1007/3-540-39967-4_13.

Baker, Tom. "DCMI Usage Board review of application profiles." Ultimo accesso 23 Gennaio 2024. <https://www.dublincore.org/specifications/dublin-core/application-profile-review/>.

Banerjee, Sourav. *Indian Food Images Dataset*, Ultimo accesso 23 Gennaio 2024, <https://www.kaggle.com/datasets/iamsouravbanerjee/indian-food-images-dataset>.

Bellavia, Fabio, Carlo Colombo, Luca Morelli, and Fabio Remondino. "Challenges in Image Matching for Cultural Heritage: An Overview and Perspective." *Image Analysis and Processing ICIAP 2022 Workshops*, (2022). DOI: 10.1007/978-3-031-13321-3.

Bennet, Mike, and Giacomo Marchioro. *iiif-prezi3 v. 1.2.1* (2023). <https://github.com/iiif-prezi/iiif-prezi3>.

Bezombes, Patrick. "JTC 21 Preliminary views on ai act timeline and horizontal/vertical integration." <https://www.eurocontrol.int/sites/default/files/2023-05/20230420-flyai-forum-session3-cen-celecec-jtc-21-preliminary-views.pdf>

Bianco, Simone, Remi Cadene, Luigi Celona, and Paolo Napoletano. "Benchmark Analysis of Representative Deep Neural Network Architectures." (2018). arXiv:1810.00736.

Biblissima. "Collections". Ultimo accesso 23 Gennaio 2024. <https://iiif.biblissima.fr/collections/about>.

Bird, Steven, Loper Edward, and Klein Ewan, *Natural Language Processing with Python*. O'Reilly Media Inc.: 2009. <https://www.nltk.org/>.

Bodleian Library. "Collection guides." Ultimo accesso 23 Gennaio 2024. <https://www.bl.uk/collection-guides/iiif>.

Bodleian Library. "Modern manuscripts and archives." Ultimo accesso 23 Gennaio 2024. <https://digital.bodleian.ox.ac.uk/collections/modern-manuscripts-and-archives/?rows=100&page=7>.

British Library. "Endangered Archives Programme." Ultimo accesso 23 Gennaio 2024. <https://eap.bl.uk/archive-file/EAP427-2-9>.

Bui, Hu, Jiayu Du, Xingyu Na, Bengu Wu, and Hao Zheng. "AISHELL-1: An open-source mandarin speech corpus and a speech recognition baseline." (2017). <https://arxiv.org/abs/1709.05522>.

Bukit Brown Working Committee. "Bukit Brown project." Ultimo accesso 23 Gennaio 2024. <https://www.bukitbrown.info/about.php>.

Cantaloupe project. "Cantaloupe." Ultimo accesso 23 Gennaio 2024. <https://cantaloupe-project.github.io/>.

- Choi, Yoonjo, Yun-Jung Yang, and Hong-Gyoo Sohn. "Resilient cultural heritage through digital cultural heritage cube: Two cases in South Korea." *Journal of Cultural Heritage* 48, (2021). DOI: 10.1016/j.culher.2021.01.007.
- Chollet, François. "Xception: Deep Learning with Depthwise Separable Convolutions." (2016). <https://arxiv.org/abs/1610.02357>.
- Cimiano, Philipp, Alexander Mädche, Steffen Staab, and Johanna Völker. "Ontology learning," in *Handbook on Ontologies*, a cura di Steffen, Staab and Rudi Studer, 245-267. Berlin: Springer, 2009.
- Cimiano, Philipp, Alexander Mädche, Steffen Staab, and Johanna Völker. "Ontology learning." in *Handbook on Ontologies*, ed. Staab Steffen, and Rudi Studer. Berlin: Springer, 2004. 245-267. DOI: 10.1007/978-3-540-24750-0
- Cimiano, Philipp. *Ontology Learning and Population from Text: Algorithms, Evaluation and Application*. New York: Springer-Verlag, 2006. DOI: 10.1007/978-0-387-39252-3.
- Çinar, Necip, Kaya Buket, and Kaya Mehemet. "Comparison of deep learning models for brain tumor classification using MRI images." *2022 International Conference on Decision Aid Sciences and Applications (DASA)*, (2022). DOI: 10.1109/DASA54658.2022.9765250.
- Cintas, Celia, Manuel Lucena, José Manuel Fuertes, Claudio Delrieux, Pablo Navarro, Rolando González-José, and Manuel Molinos. "Automatic feature extraction and classification of Iberian ceramics based on deep convolutional networks." *Journal of Cultural Heritage* 41, (2020). DOI: 10.1016/j.culher.2019.06.005.
- Cliché, Mathieu, David Rosenberg, Dhruv Madeka, and Connie Yee. "Scatteract: Automated extraction of data from scatter plots." (2017). <https://arxiv.org/abs/1704.06687>.
- COCO Dataset. Ultimo accesso 23 Gennaio 2024. <https://cocodataset.org/#home>.
- Commissione Europea, Expert Group on Digital Cultural Heritage, and Europeana. *Basic principles and tips for 3D digitisation of tangible cultural heritage for cultural heritage professionals and institutions and other custodians of cultural heritage* (2020), <https://digital-strategy.ec.europa.eu/en/library/basic-principles-and-tips-3d-digitisation-cultural-heritage>.
- Commissione Europea. "Lotta contro il traffico illecito di beni culturali." Ultimo accesso 23 Gennaio 2024. <https://culture.ec.europa.eu/it/cultural-heritage/cultural-heritage-in-eu-policies/protection-against-illicit-trafficking>. Ultimo accesso 17 Gennaio 2024.
- Commissione Europea. *2016/C 202/389 Charter of Fundamental Rights of the European Union* (2016). <https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:12016P/TXT>.
- Commissione Europea. *Artificial Intelligence Act* (2023), [https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI\(2021\)698792_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf)
- Commissione Europea. *COM/2021/206 Proposta di Regolamento del parlamento europeo e del consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (legge sull'intelligenza artificiale) e modifica alcuni atti legislativi dell'unione* (2021). https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0006.02/DOC_1&format=PDF.
- Commissione Europea. *Draft standardisation request to the European Standardisation Organisations in support of safe and trustworthy artificial intelligence* (2022). <https://ec.europa.eu/docsroom/documents/52376>.

- Committee on Culture and Education. *AI in Education, Culture and the Audiovisual sector* (Parlamento Europeo 2021). https://www.europarl.europa.eu/doceo/document/TA-9-2021-0238_IT.pdf.
- Condorelli, Francesca, Fulvio Rinaudo, Francesco Salvatore and Stefano Tagliaventi. "A Neural Networks Approach to Detecting Lost Heritage in Historical Video." *International Journal of Geo-Information* 9, n. 5 (2020). DOI: 10.3390/ijgi9050297.
- Consiglio d'Europa, "Organizzazione delle Nazioni Unite per l'Educazione, la Scienza e la Cultura (UNESCO)," <https://coe.int/it/web/cultural-routes/unesco>. Ultimo accesso 12 Gennaio 2024.
- Consorzio Interuniversitario AlmaLaurea, *XXIII Indagine Profilo dei Laureati 2020. Rapporto 2021* (2021), https://www.almalaurea.it/sites/default/files/2022-05/almalaurea_profilo_rapporto2021_0.pdf.
- Coşkun, Musab, Özal Yildirim, Ayşegül Uçar, and Yakup Demir. "An Overview of Popular Deep Learning Methods." *European Journal of Technique (EJT)* 7, n. 2 (2017). <https://dergipark.org.tr/en/download/article-file/437659>.
- Costituzione della Repubblica Italiana*. <https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:costituzione:1947-12-27~art9>.
- D. Lgs. 42/2004 *Codice dei beni culturali e del paesaggio* (2004). <https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:decreto.legislativo:2004-01-22;42>.
- D.Lgs. 82/2005 *Codice dell'Amministrazione Digitale* (2005). <https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:decreto.legislativo:2005-03-07;82>.
- Darapaneni, Narayana, Krishnamurthy B, and Anwesh Reddy Paduri. "Convolution Neural Networks: A Comparative Study for Image Classification." *15th (IEEE) International Conference on Industrial and Information Systems (ICIIS)*, (2020). DOI: 10.1109/ICIIS51140.2020.9342667.
- Davis-Marks, Isis. "Lost Edges of Rembrandt's 'Night Watch' Are Restored Using Artificial Intelligence." *Smithsonian Magazine*, 25 Giugno 2021. <https://www.smithsonianmag.com/smart-news/lost-edges-rembrandts-night-watch-are-restored-using-artificial-intelligence-180978056/>.
- DBpedia. "Semantic Web Stack." Ultimo accesso 23 Gennaio 2024. https://dbpedia.org/page/Semantic_Web_Stack.
- De Vincentis, Stefania, and Martin Critelli. "Mappare il museo in IIF : Una combinazione di deep zoom e VR360 per la Galleria Borghese di Roma." *La memoria digitale: forme del testo e organizzazione della conoscenza. Atti del XII Convegno Annuale AIUCD*, a cura di Carbé, Emmanuela, Gabriele Lo Piccolo, Alessia Valenti, e Francesco Stella. Siena. ISBN – 978-88-942535-7-3.
- Degu Abate, A. Agapiou, K. Toumbas, A. Lampropoulos, K. Petrides, R. Pierdicca, M. Paolanti, F. Di Stefano, A. Felicetti, and E. S. Malinverni. "Artificial Intelligence to Fight Illicit Trafficking of Cultural Property," *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 48M2, (2023). DOI: 10.5194/isprs-archives-XLVIII-M-2-2023-3-2023.
- Della Piazza, Sara. "Il caso Cambridge Analytica." *DirittoConsenso*, 21 Dicembre 2021, <https://www.dirittoconsenso.it/2021/12/21/il-caso-cambridge-analytica/>.
- Deng, Jia, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. "Imagenet: A large-scale hierarchical image database." *2009 IEEE conference on computer vision and pattern recognition*, (2009). DOI: 10.1109/CVPR.2009.5206848.

- Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Tautanova. "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding." *Proceedings of NAACL-HLT*, (2019). <https://github.com/google-research/bert>. arXiv:1810.04805v2.
- Dhillon, Anamika, and Gyanendra K. Verma. "Convolutional neural network: a review of models, methodologies and applications to object detection." *Progress in Artificial Intelligence* 9, (2020). DOI: 10.1007/s13748-019-00203-0.
- Digital Vatican Library. "La Biblioteca Apostolica Vaticana." Ultimo accesso 23 Gennaio 2024. <https://spotlight.vatlib.it/overview>.
- Du, Xing, Yongfu Sun, Yupeng Song, Huifeng Sun, and Lei Yang. "A Comparative Study of Different CNN Models and Transfer Learning Effect for Underwater Object Classification in Side-Scan Sonar Images." *Remote Sensing* 15, n. 3: 593. DOI: 10.3390/rs15030593.
- EBU Operating Eurovision and Euroradio. *TECH 3293 EBU CORE METADATA SET (EBUCore) Specification v. 1.10. 2020*. <https://tech.ebu.ch/docs/tech/tech3293.pdf>.
- Enciclopedia Treccani. "Pitch." Ultimo accesso 23 Gennaio 2024. www.treccani.it/enciclopedia/pitch/.
- Enciclopedia Treccani. "Query." Ultimo accesso 23 Gennaio 2024. www.treccani.it/vocabolario/query/.
- ENIGMA. "ENIGMA". Ultimo accesso 23 Gennaio 2024. <https://eu-enigma.eu/work-packages/>.
- Europeana Cloud. *D4.4 Recommendations for enhancing EDM to represent digital Content* (2016), https://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/Europeana_Cloud/Deliverables/D4.4%20Recommendations%20For%20Enhancing%20EDM%20to%20Support%20Research%20Oriented%20Content.pdf
- EUROSTAT. "Enterprises with an increase in the remote access to their ICT systems due to the Covid-19, by activity, EU 2020." Ultimo accesso 23 Gennaio 2024. [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=File:Enterprises with an increase in the remote access to their ICT systems due to the Covid-19, by activity, EU 2020.png](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=File:Enterprises_with_an_increase_in_the_remote_access_to_their_ict_systems_due_to_the_covid-19_by_activity_eu_2020.png)
- EUROSTAT. "Government expenditure on cultural, broadcasting and publishing services." Ultimo accesso 23 Gennaio 2024. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Government_expenditure_on_cultural_broadcasting_and_publishing_services#General_government_expenditure_on_cultural_services.2C_broadcasting_and_publishing_services.
- Everingham, Mark, Ali S. M. Eslami, Luc Van Gool, Christopher K. I. Williams, John Winn, and Andrew Zisserman. "The PASCAL Visual Object Classes Challenge: A Retrospective." *International Journal of Computer Vision* 111, n. 1 (2015). DOI 10.1007/s11263-014-0733-5.
- EXPLOSION. "spaCy." Ultimo accesso 23 Gennaio 2024. <https://spacy.io/>.
- Ferraino, Giuliana. "Il Forum di Davos, in diretta." *Corriere della Sera*, 18 Gennaio 2024, https://www.corriere.it/economia/finanza/diretta-live/24_gennaio_18/forum-davos-diretta-herzog-prima-una-soluzione-politica-vogliamo-certezza-pace-d0ee8932-b5cf-11ee-8131-eee5ee6b3690.shtml?refresh_ce.
- Fiorucci, Marco, Marina Khoroshiltseva, Massimiliano Pontil, Arianna Traviglia, Alessio Del Bue, and Stuart James. "Machine Learning for Cultural Heritage: A Survey." *Pattern Recognition Letters* 133, (2020). DOI: 10.1013/j.patrec.2020.02.017.
- Fisher, Yu, Haofeng Chen, Xin Wang, Wenqi Xian, Yingying Chen, Fangchen Liu, Vashisht Madhavan, and Trevor Darrell. "BDD100K : A Diverse Driving Dataset for Heterogeneous Multitask Learning." 2018. <https://arxiv.org/abs/1805.04687>.

- Floridi, Luciano, and Josh Cowl. "A Unified Framework of Five Principles for Ethical AI." *Harvard Data Science Review* 1, n. 1 (2019). DOI: 10.1162/99608f92.8cd550d1.
- Floridi, Luciano. "Translating Principles into Practices of Digital Ethics: Five Risks of Being Unethical." *Philosophy & Technology* 32, (2019). DOI: 10.1007/s13347-019-00354-x.
- FlyerZone, "Yorescape." Ultimo accesso 23 Gennaio 2024. <https://yorescape.com/>.
- Food and Drug Administration (FDA). "Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices." Ultima modifica 19 Ottobre 2023. <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices>.
- Fragrasso, Beatrice. "La responsabilità penale del produttore di sistemi di intelligenza artificiale." *Diritto Penale Contemporaneo. Rivista Trimestrale*, 1/2023, (2023). https://www.sistemapenale.it/pdf_contenuti/1686584973_fragasso-la-responsabilita-penale-del-produttore-di-sistemi-di-ia.pdf.
- Franchina, Luisa. "Dichiarazione di Bletchley, svolta per lo sviluppo responsabile dell'intelligenza artificiale: ecco perché." *Cybersecurity 360*, 8 Novembre 2023, <https://www.cybersecurity360.it/cultura-cyber/dichiarazione-di-bletchley-svolta-per-lo-sviluppo-responsabile-dellintelligenza-artificiale-ecco-perche/>.
- Freire, Nuno. "The Data Aggregation Lab Software. Experimentation for Linked Data Aggregation in Cultural Heritage." *Metadata and Semantic Research. MTSR 2019. Communications in Computer and Information Science* 1057, (2019). DOI: 10.1007/978-3-030-36599-8_38.
- Garante per la protezione dei dati personali. "AI – Intelligenza artificiale." Ultimo accesso 23 Gennaio 2024. <https://www.garanteprivacy.it/temi/intelligenza-artificiale>.
- Garante per la protezione dei dati personali. "Intelligenza artificiale: Garante privacy apre un'indagine sulla raccolta di dati personali on line per addestrare gli algoritmi. L'iniziativa è volta a verificare l'adozione di misure di sicurezza da parte di siti pubblici e privati." Ultimo accesso 23 Gennaio 2024. <https://www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/9952078>.
- Garante per la protezione dei dati personali. *Decalogo per la realizzazione di servizi sanitari nazionali attraverso sistemi di Intelligenza Artificiale* (2021). <https://www.garanteprivacy.it/documents/10160/0/Decalogo+per+la+realizzazione+di+servizi+sanitari+nazionali+attraverso+sistemi+di+Intelligenza+Artificiale.pdf/a5c4a24d-4823-e014-93bf-1543f1331670?version=2.0&LinkSource=PassleApp>.
- Geiger, Andreas, Philip Lenz, Christoph Stiller, and Raquel Urtasun. "Vision meets Robotics: The KITTI Dataset. *International Journal of Robotics Research (IJRR)*". 2013. <https://www.cvlibs.net/publications/Geiger2013IJRR.pdf>
- Getty Institute. "Metadata Standards Crosswalk." Ultimo accesso 23 Gennaio 2024. https://www.getty.edu/research/publications/electronic_publications/intrometadata/crosswalks.html.
- Getty Museum. "Collection." Ultimo accesso 23 Gennaio 2024. <https://www.getty.edu/art/collection/mirador/?manifest=https://media.getty.edu/iiif/manifest/53be857e-41e8-4198-b45d-2e0f52d3051b>.
- Girshick, Ross, Jeff Donahue, Trevor Darrell, and Jitendra Malik. "Rich feature hierarchies for accurate object detection and semantic segmentation." *Computer Vision and Pattern Recognition*, (2014). <https://arxiv.org/abs/1311.2524>.
- Girshick, Ross. "Fast r-cnn." (2015). <https://arxiv.org/abs/1504.08083>.

- GNU. "Libextractor". Ultimo accesso 23 Gennaio 2024. <https://www.gnu.org/software/libextractor/>.
- Godfrey, John J., and Edward Holliman. "Switchboard-1 Release 2 LDC97S62." (1993). DOI:10.35111/sw3hrw02.
- Google Cloud. "Guida avanzata a Inception v3." Ultimo accesso 23 Gennaio 2024. <https://cloud.google.com/tpu/docs/inception-v3-advanced?hl=it>.
- Google. "Open Images v6." Ultimo accesso 23 Gennaio 2024. <https://storage.googleapis.com/openimages/web/index.html>.
- Google. "Open Images v7." Ultimo accesso 23 Gennaio 2024. https://storage.googleapis.com/openimages/web/factsfigures_v7.html.
- Google. "Speech-to-Text." Ultimo accesso 23 Gennaio 2024. <https://cloud.google.com/speech-to-text/docs/basics?hl=it>.
- Goyal, Archana, Vishal Gupta, and Manish Kumar. "Recent Named Entity Recognition and Classification techniques: A systematic review." *Computer Science Review* 29, (2018). DOI: 10.1016/j.cosrev.2018.06.001.
- Greco Gianluigi. "L'IA e le sue applicazioni: soluzioni per la complessità." *Lezione al Master di II Livello in Conservatore dei Documenti Digitali* (20 Gennaio 2024), Università della Calabria.
- Grilli, Eleonora, Elisa Mariarosaria Farella, Alessandro Torresani, and Fabio Remondino. "Geometric Features Analysis for the Classification of Cultural Heritage Point Clouds." *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-2/W15*, (2019). DOI: 10.5194/isprs-archives-XLII-2-W15-541-2019.
- Grootendorst, Maarten. "BERTopic: Neural topic modeling with a class-based TF-IDF procedure." (2022). <https://arxiv.org/pdf/2203.05794.pdf>.
- Guerrini, Mauro, and Tiziana Possemato. *Linked Data per biblioteche, archivi e musei*. Milano: Editrice Bibliografica, 2015. ISBD: 978-8870758306.
- Guerrini, Mauro. *Metadattazione*. Milano: Editrice Bibliografica, 2022. ISBD: 978-8893575195.
- Han, Kyu J., Ramon Prieto, Kaixing Wu, and Tao Ma. "State-of-the-Art Speech Recognition Using Multi-Stream Self-Attention With Dilated 1D Convolutions." (2019). <https://arxiv.org/abs/1910.00716>.
- Haridasa, Arul Valiyavalappil, Ramalatha Marimuthub, and Vaazi Gangadharan Sivakumarc. "A critical review and analysis on techniques of speech recognition: The road ahead." *International Journal of Knowledge-based and Intelligent Engineering Systems* 22, n. 1 (2018). DOI: 10.3233/KES-180374.
- Harvard Library. "File Information Tool Set (FITS)." Ultimo accesso 23 Gennaio 2024. <https://projects.iq.harvard.edu/fits/fits-processing>.
- Harvey, Phil. "ExifTool." Ultimo accesso 23 Gennaio 2024. <https://exiftool.org/>.
- He, Kaiming, Georgia Gkioxari, Piotr Dollar, and Ross Girshick. "Mask R-CNN." (2018). arXiv:1703.06870.
- He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep Residual Learning for Image Recognition." (2015). <https://arxiv.org/abs/1512.03385>.
- Hutchinson, Tim. "Protecting privacy in the archives: Preliminary explorations of topic modeling for born-digital collections." *2017 IEEE International Conference on Big Data (Big Data)*, (2018). DOI:10.1109/BigData.2017.8258177

- Hutchinson, Tim. "Protecting privacy in the archives: Preliminary explorations of topic modeling for born-digital collections." *2017 IEEE International Conference on Big Data (Big Data)* 2251-55, (2018). DOI:10.1109/BigData.2017.8258177.
- ICOM International Committee for Documentation Special Interest Group. *Conceptual Reference Model (CIDOC CRM) v. 7.1.2* (2022). https://www.cidoc-crm.org/sites/default/files/cidoc_crm_v7.1.2.pdf.
- IIP. "IIPImage Server." <http://iipimage.sourceforge.net/documentation/server/>. Ultimo accesso Maggio 2023.
- ImageNet. "Large Scale Visual Recognition Challenge." Ultimo accesso 23 Gennaio 2024. <https://www.image-net.org/challenges/LSVRC/>.
- Imran, Saqib, Rizwan Ali Naqvi, Muhammad Sajid, Tauqeer Safdar Malik, Saif Ullah, Syed Atif Moqurrab, and Dong Keon Yon. "Artistic Style Recognition: Combining Deep and Shallow Neural Networks for Painting Classification." *Mathematics* 11, n. 22 (2023): 4564. DOI: 10.3390/math11224564.
- International Association of Sound and Audiovisual Archives (IASA). *Guidelines on the Production and Preservation of Digital Audio Objects, IASA-TC04, Second Edition* (2009). <https://www.iasa-web.org/tc04/audio-preservation>.
- International Committee for Documentation (CIDOC) Conceptual Reference Model (CRM). "Versions of CIDOC-CRM." Ultimo accesso 27 Gennaio 2024. <https://www.cidoc-crm.org/versions-of-the-cidoc-crm>.
- International Committee for Documentation (CIDOC), and International Council of Museums (ICOM). *Conceptual Reference Model (CRM) v. 7.1.2* (2022). https://www.cidoc-crm.org/sites/default/files/cidoc_crm_v7.1.2.pdf.
- International Council on Archives. *ISAD (G): General International Standards Archival Description – Second edition* (2000). <https://www.ica.org/en/isadg-general-international-standard-archival-description-second-edition>.
- International Federation of Library Associations and Institutions - Rare Book and Special Collections Section. *Linee guida per pianificare la digitalizzazione di collezioni di libri rari e manoscritti* (2014). <https://repository.ifla.org/bitstream/123456789/458/1/guidelines-for-planning-digitization-it.pdf>.
- International Federation of Library Associations and Institutions. *ISBD: International Standard Bibliographic Description. Consolidated edition*. Edizione italiana a cura dell'ICCU, 2012. <https://www.iccu.sbn.it/it/eventi-novita/novita/ISBD-International-Standard-Bibliographic-Description.-Consolidated-edition.-Edizione-italiana-novembre-2012/>.
- International Image Interoperability Framework (IIIF). "API cookbook." Ultimo accesso 23 Gennaio 2024. <https://iiif.io/api/cookbook/>.
- International Image Interoperability Framework (IIIF). "API Specifications." Ultimo accesso 23 Gennaio 2024. <https://iiif.io/api/>.
- International Image Interoperability Framework (IIIF). "Image 3.0 Change Log." Ultimo accesso 23 Gennaio 2024. <https://iiif.io/api/image/3.0/change-log/>,
- International Image Interoperability Framework (IIIF). "Presentation API 3.0." Ultimo accesso 23 Gennaio 2024. <https://iiif.io/api/presentation/3.0/>.
- International Research on Permanent Authentic Records in Electronic System (InterPARES). "The InterPARES Project." Ultimo accesso 23 Gennaio 2024. <http://www.interpares.org/index.htm>
- Invenio. "flask-iiif." Ultimo accesso 23 Gennaio 2024. <https://github.com/inveniosoftware/flask-iiif>.

- Iosifova, Olena, Ievgen Iosifov, Volodymyr Sokolov, Oleh Romanovskyi, and Igor Sukaylo. "Analysis of Automatic Speech Recognition Methods." *CEUR Workshop Proceedings* 3175, (2021). <https://ceur-ws.org/Vol-2923/paper27.pdf>
- Isaac, Antoine. "The state of IIF at Europeana." *OCLC Research Mini-symposium on the Discovery and Use of Open Collections*, (2019). <https://www.oclc.org/content/dam/research/presentations/2019/061919-Isaac.pdf>.
- Ismail, Safaa S. I., Romany F. Mansour, Rasha M. Abd El-Aziz, and Ahmed I. Talobaet. "Efficient E-Mail Spam Detection Strategy Using Genetic Decision Tree Processing with NLP Features." *Computational Intelligence and Neuroscience*, (2022). Doi: 10.1155/2022/7710005.
- ISO 15836-2003. *Information and documentation - The Dublin Core™ metadata element set* (2003). <https://www.dublincore.org/specifications/dublin-core/dcmi-terms/>.
- ISO 639-1:2002. *Codes for the representation of names of languages – Part 1 : Alpha-2 code*.
- Istituto Centrale del Catalogo Unico delle biblioteche e per le informazioni bibliografiche. "CulturalItalia." Ultimo accesso 23 Gennaio 2024. <https://www.culturaitalia.it/>.
- Istituto Centrale del Catalogo Unico delle biblioteche e per le informazioni bibliografiche. "Internet Culturale." Ultimo accesso 23 Gennaio 2024. <https://www.internetculturale.it/>.
- Istituto Centrale per il Catalogo e la Documentazione (ICCD). "Schede di catalogo." Ultimo accesso 23 Gennaio 2024. <https://github.com/ICCD-MiBACT/Standard-catalografici/tree/master/schede-di-catalogo>.
- Istituto Centrale per la Digitalizzazione del Patrimonio Culturale – Digital Library. *Linee guida per i processi di digitalizzazione 2022-2023* (2023). <https://docs.italia.it/italia/icdp/icdp-pnd-digitalizzazione-docs/it/v1.0-giugno-2022/il-progetto-di-digitalizzazione/come.html>.
- Istituto Centrale per la Digitalizzazione del Patrimonio Culturale – Digital Library. "Il Piano." Ultimo accesso 23 Gennaio 2024. <https://digitallibrary.cultura.gov.it/il-piano/>.
- Istituto Centrale per la Digitalizzazione del Patrimonio Culturale – Digital Library. "Linee guida." Ultimo accesso 23 Gennaio 2024. <https://digitallibrary.cultura.gov.it/linee-guida/>.
- Istituto Centrale per la Digitalizzazione del Patrimonio Culturale – Digital Library. *Linee guida per i processi di digitalizzazione 2022-2023* (2022). <https://docs.italia.it/italia/icdp/icdp-pnd-digitalizzazione-docs/it/v1.0-giugno-2022/index.html>.
- Istituto Centrale per la Digitalizzazione del Patrimonio Culturale – Digital Library. *Linee guida per la redazione del piano di gestione dei dati* (2022). <https://docs.italia.it/italia/icdp/icdp-pnd-dmp-docs/it/v1.0-giugno-2022/index.html>.
- Istituto Centrale per la Digitalizzazione del Patrimonio Culturale – Digital Library. *Linee guida per l'acquisizione, la circolazione e il riuso delle riproduzioni dei beni culturali in ambiente digitale* (2022). <https://docs.italia.it/italia/icdp/icdp-pnd-circolazione-riuso-docs/it/v1.0-giugno-2022/index.html>.
- Istituto Centrale per la Digitalizzazione del Patrimonio Culturale – Digital Library. *Linee guida per la classificazione di prodotti e servizi digitali, processi e modelli di gestione* (2022). <https://docs.italia.it/italia/icdp/icdp-pnd-servizi-docs/it/v1.0-giugno-2022/index.html>.
- Istituto Centrale per la Digitalizzazione del Patrimonio Culturale – Digital Library. *Introduzione alla metodologia per la valutazione della maturità digitale degli istituti culturali* (2022). <https://docs.italia.it/italia/icdp/icdp-pnd-maturita-docs/it/v1.0-giugno-2022/index.html>.

- Istituto Centrale per la Digitalizzazione del Patrimonio Culturale – Digital Library. *Piano nazionale per la digitalizzazione del patrimonio culturale 2022-2023* (2022). https://digitallibrary.cultura.gov.it/wp-content/uploads/2023/04/PND_V1_1_2023_v2.pdf.
- Istituto Centrale per la Digitalizzazione del Patrimonio Culturale – Digital Library. “MICROFILM: al via la digitalizzazione dei manoscritti.” Ultimo accesso 27 Gennaio 2024. <https://pnrr.cultura.gov.it/microfilm-al-via-la-digitalizzazione-dei-manoscritti-m1c31-1-strategie-e-piattaforme-digitali-per-il-patrimonio-culturale/>.
- Istituto Centrale per la Digitalizzazione del Patrimonio Culturale – Digital Library. “PNRR, digitalizzazione del patrimonio culturale: online la procedura di gara da 27,9 milioni di euro per la categoria ‘Carta: documentazione catastale (registri e mappe) e giornali quotidiani postunitari’.” Ultimo accesso 27 Gennaio 2024. <https://digitallibrary.cultura.gov.it/avvisi/pnrr-digitalizzazione-del-patrimonio-culturale-online-la-procedura-di-gara-da-279-milioni-di-euro-per-la-categoria-carta-documentazione-catastale-registri-e-mappe-e-giornali-quotidiani/>.
- Istituto Centrale per la Digitalizzazione del Patrimonio Culturale – Digital Library. “PNRR, digitalizzazione del patrimonio culturale: online la procedura di gara da 16,9 milioni di euro per la categoria ‘Archivi fotografici: positivi, negativi, unicum e disegni’.” Ultimo accesso 27 Gennaio 2024. <https://digitallibrary.cultura.gov.it/notizie/pnrr-digitalizzazione-del-patrimonio-culturale-online-la-procedura-di-gara-da-169-milioni-di-euro-per-la-categoria-archivi-fotografici-positivi-negativi-unicum-e-disegni/>.
- Istituto Centrale per la Digitalizzazione del Patrimonio Culturale – Digital Library. “PNRR, digitalizzazione del patrimonio culturale: online la procedura di gara da 18,6 milioni di euro per la categoria ‘Oggetti museali: beni di deposito (storico-artistici, archeologici) e grafici’.” Ultimo accesso 27 Gennaio 2024. <https://digitallibrary.cultura.gov.it/notizie/pnrr-digitalizzazione-del-patrimonio-culturale-online-la-procedura-di-gara-da-186-milioni-di-euro-per-la-categoria-oggetti-museali-beni-di-deposito-storico-artistici-archeologici-e-gra/>.
- Istituto Nazionale Ricerche Turistiche (ISNART). “Report indagine sul turismo nazionale 2022.” Ultimo accesso 23 Gennaio 2024. <https://www.isnart.it/it/economia-del-turismo/indagini-nazionali/indagine-2022/>.
- Jas, Ram. *UNESCO Heritage sites* (2021). <https://www.kaggle.com/datasets/ramjasmaurya/unesco-heritage-sites2021>
- Java Native Access. “MediaInfo.” Ultimo accesso 23 Gennaio 2024. <https://github.com/java-native-access/jna>.
- Ježek Elisabetta, and Rachele Sprugnoli. *Linguistica computazionale. Introduzione all’analisi automatica dei testi*. Il Mulino, 2023. ISBD: 978-8815290359.
- Jocher, Glenn. “Ultralytics.” Ultimo accesso 23 Gennaio 2024. <https://github.com/ultralytics/ultralytics>.
- Jocher, Glenn. “Ultralytics.” Ultimo accesso 23 Gennaio 2024. <https://ultralytics.com/yolov8>.
- Jurafsky, Daniel, and James H. Martin. *Speech and Language Processing An Introduction to Natural Language Processing. Computational Linguistics, and Speech Recognition*. 2023. <https://web.stanford.edu/~jurafsky/slp3/ed3book.pdf>
- Kahneman, Daniel. *Thinking Fast and Slow*. United Kingdom: Penguin Books, 2012. ISBN: 978-0141033570.

- Kavitha Srinivasan, S. Mohanavalli, B. Bharathi, C. H. Rahul, S. Shailesh & K. Preethi. "Classification of Indian Monument Architecture Styles Using Bi-Level Hybrid Learning Techniques." *Inventive Systems and Control. Lecture Notes in Networks and Systems* 436, (2023). DOI: 10.1007/978-981-19-1012-8_32.
- Chollet, François. "API Applications." Ultimo accesso 23 Gennaio 2024. <https://keras.io/api/applications/>.
- Chollet, François. "Keras." Ultimo accesso 23 Gennaio 2024. <https://keras.io/>.
- Chollet, François. "Video Classification." Ultimo accesso 23 Gennaio 2024. https://keras.io/examples/vision/video_classification/.
- Khanal, Bimarsha. *Nepali Cultural Dress and Ornaments*. Ultimo accesso 23 Gennaio 2024. <https://www.kaggle.com/datasets/bimarshakhanal/nepali-cultural-dress-and-ornaments>.
- Kobzev, Ivan, and Vasiliev Roman. *Architectural Heritage Elements Image64 Dataset* (2022). <https://www.kaggle.com/datasets/ikobzev/architectural-heritage-elements-image64-dataset>.
- Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "ImageNet Classification with Deep Convolutional Neural Networks." *Advances in Neural Information Processing Systems 25 (NIPS 2012)*, (2012). https://papers.nips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf.
- Kumar, Pakhee, Ferda Ofli, Muhammad Imran, and Carlos Castillo. "Detection of Disaster-Affected Cultural Heritage Sites from Social Media Images Using Deep Learning Techniques." *Journal on Computational and Cultural Heritage* 13 (2020). DOI: 10.1145/3383314.
- Kyrlitsias, Christos, Maria Christofi, Despina Michael-Grigoriou, Domna Banakou, and Andri Ioannou. "Corrigendum: A Virtual Tour of a Hardly Accessible Archaeological Site: The Effect of Immersive Virtual Reality in User Experience, Learning and Attitude Change." *Frontiers in Computer Science* 3, (2021). DOI: 10.3389/fcomp.2021.697259.
- Lakshmanan, Valliappa, Martin Görner, and Ryan Gillard. *Practical Machine Learning for Computer Vision*. O'Reilly Media Inc. 2021. ISBN: 9781098102364.
- Lan Zhenzhong, Mingda Chen, Sebastian Goodman, Kevin Gimpel, Piyush Sharma, and Radu Soricutet. "ALBERT: A Lite BERT for Self-supervised Learning of Language Représentations." (2020). <https://arxiv.org/abs/1909.11942>.
- Lazzaretti, Luciana. "What is the role of culture facing the digital revolution challenge? Some reflections for a research agenda." *European Planning Studies* 30, n. 9 (2022). DOI: 10.1080/09654313.2020.1836133
- Li, Yue, Xutao Wang, and Pengjian Xu. "Chinese Text Classification Model Based on Deep Learning." *Future Internet* 10, n. 11 (2018): 113. DOI: 10.3390/fi10110113.
- Library of Congress. "An introduction to VRA Core 4." Ultimo accesso 23 Gennaio 2024. https://www.loc.gov/standards/vracore/VRA_Core4_Intro.pdf.
- Library of Congress. "DC cross." Ultimo accesso 23 Gennaio 2024. <http://lcweb.loc.gov/marc/dccross.html>.
- Library of Congress. *Metadata Encoding & Transmission Standard v. 1.12.1*. Ultimo accesso 23 Gennaio 2024. <https://www.loc.gov/standards/mets/mets-schemadocs.html>.
- Library of Congress. *PREMIS Data Dictionary for Preservation Metadata v.3.0* (2015). <https://www.loc.gov/standards/premis/v3/premis-3-0-final.pdf>.
- Library of Congress. *Visual Resources Association Core (VRA Core) v.4* (2007). <https://www.loc.gov/standards/vracore/schemas.html>.

- Lin, Tsung-Yi, Michael Maire, Serge Belongie, Lubomir Bourdev, Ross Girshick, James Hays, Pietro Perona, Deva Ramanan, C. Lawrence Zitnick, and Piotr Dollár. "Microsoft COCO: Common Objects in Context." (2015). <https://arxiv.org/abs/1405.0312>.
- Liu Yinhan, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. "RoBERTa : A Robustly Optimized BERT Pretraining Approach," (2019). arxiv.org/pdf/1907.11692.
- Liu, Xiaoyi, Diego Klabjan, and Patrick N. Bless. "Data extraction from Charts via Single Deep Neural Network." (2019). <https://arxiv.org/abs/1906.11906>.
- Liu, Yuxiang, Qiaoliang Xiang, Ye Wang, and L. Cai. "Cultural style based music classification of audio signals." *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, (2009). DOI: 10.1109/ICASSP.2009.4959519.
- Lopez, Patrice. "Grobid". Ultimo accesso 23 Gennaio 2024. <https://github.com/kermitt2/grobid>.
- Loris Image Server. "Loris." Ultimo accesso 23 Gennaio 2024. <https://github.com/loris-imageserver/loris>.
- Loris Image server. "Releases." Ultimo accesso 23 Gennaio 2024. <https://github.com/loris-imageserver/loris/releases>.
- Luo, Junyu, Zekun Li, Jinpeng Wang, and Chin-Yew Lin. "ChartOCR: Data Extraction from Charts Images via a Deep Hybrid Framework." *2021 IEEE Winter Conference on Applications of Computer Vision (WACV)*, (2021). https://openaccess.thecvf.com/content/WACV2021/papers/Luo_ChartOCR_Data_Extraction_From_Charts_Images_via_a_Deep_Hybrid_WACV_2021_paper.pdf
- Luo, Junyu. "DeepRule". Ultimo accesso 23 Gennaio 2024. <https://github.com/soap117/DeepRule>.
- Mai, Chan Lois, and Marcia Zheng. "Metadata Interoperability and Standardization – A Study of Methodology, Part I: Achieving Interoperability at the Schema Level," *D-Lib Magazine* 12, n. 6 (2006). DOI:10.1045/june2006-zeng.
- Mall, Pawan Kumar, Pradeep Kumar Singh, Swapnita Srivastav, Vipul Narayan, Marcin Paprzycki, Tatiana Jaworska, and Maria Ganzha. "A comprehensive review of deep neural networks for medical image processing: Recent developments and future opportunities." *Healthcare Analytics* 4, (2023). DOI: 10.1016/j.health.2023.100216.
- Manning, Christopher D., Mihai Surdeanu, John Bauer, Jenny Finkel, Steven J. Bethard, and David McClosky. "The Stanford CoreNLP Natural Language Processing Toolkit." *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, (2014). DOI: 10.3115/v1/P14-5010.
- Manning, Christopher D., Prabhakar Raghavan, and Hinrich Schütze. *Introduction to Information Retrieval*. Cambridge University Press Online edition, 2009. <https://nlp.stanford.edu/IR-book/pdf/irbookonlinereading.pdf>.
- Manning, Christopher D., Prabhakar Raghavan, and Hinrich Schütze. *Introduction to Information Retrieval*. Cambridge University Press. Online edition, 2009. <https://nlp.stanford.edu/IR-book/pdf/irbookonlinereading.pdf>
- Marinescu, Maria-Cristina, Artem Reshetnikov, and Joaquim More López. "Improving object detection in paintings based on time contexts," *International Conference on Data Mining Workshops (ICDMW)*, (2020). DOI: 10.1109/ICDMW51313.2020.00133.

Masci, Maria Emilia. *VRA Core 4.0 Descrizione degli Elementi – Traduzione italiana*. Scuola Normale Superiore di Pisa, 2009.
https://www.loc.gov/standards/vracore/VRA_Core4_Element_Description_ITA.pdf.

McCallum, Andrew. “Metatagger.” Ultimo accesso 23 Gennaio 2024. <https://github.com/iesl/rexa1-metatagger>.

McGath, Gary. “Ffident.” Ultimo accesso 23 Gennaio 2024. <https://github.com/gmcgath/ffident>.

Media Library Online. “Archivio Storico Gazzetta di Parma.” Ultimo accesso 23 Gennaio 2024.
<https://archiviogazzettadiparma.medialibrary.it/home/index.aspx>.

Media Library Online. “MLOL DH per la digitalizzazione dei beni culturali.” Ultimo accesso 23 Gennaio 2024.
<https://www.medialibrary.it/pagine/pagina.aspx?id=736>.

Melo, Dora, Irene Pimenta Rodrigues, and Davide Varagnolo. “A strategy for archives metadata representation on CIDOC-CRM and knowledge discovery.” *Semantic Web 14*, (2023). DOI: 10.3233/SW-222798.

Mengmeng, Liu, and Jun Zhang. “Gesture estimation for 3D martial arts based on neural network.” *Displays 72* (2022). DOI: 10.1016/j.displa.2021.102138.

Mensink, Thomas, and Jan van Gemert. “The Rijksmuseum challenge: Museum-centered visual recognition”. *ICMR '14: Proceedings of International Conference on Multimedia Retrieval*, 2014. DOI: 10.1145/2578726.2578791.

Michon, Elise, Minh Quang Pham, Josep Crego, and Jean Senellart. “Neural Network Architectures for Arabic Dialect Identification.” *Proceedings of the Fifth Workshop on NLP for Similar Languages, Varieties and Dialects (VarDial 2018)*, (2018). <https://aclanthology.org/W18-3914.pdf>

Ministero della Cultura (MiC). “Cultura 4.0.” Ultimo accesso 23 Gennaio 2024. <https://pnrr.cultura.gov.it/>.

Ministero della Cultura (MiC). “Dal caos al cosmo.” Ultimo accesso 23 Gennaio 2024.
<https://www.beniculturali.it/evento/dal-caos-al-cosmo-la-nuova-teca-digitale-dellarchivio-centrale-dello-stato>.

Ministero della Cultura (MiC). “Dati della cultura.” Ultimo accesso 23 Gennaio 2024.
<https://www.beniculturali.it/dati-della-cultura>.

Ministero della Cultura (MiC). “Grand virtual tour.” Ultimo accesso 23 Gennaio 2024.
<https://www.beniculturali.it/virtualtour>.

Ministero della Cultura (MiC). “L’attività del Comando Carabinieri per la Tutela del Patrimonio Culturale nel 2022.” Ultimo accesso 17 Gennaio 2024.
<https://www.beniculturali.it/comunicato/24523#:~:text=Tra%20i%20recuperi%20del%202022,cui%20951%20di%20arte%20contemporanea>.

Ministero della Cultura (MiC). “Organigramma.” Ultimo accesso 23 Gennaio 2024.
https://media.beniculturali.it/mibac/files/boards/be78e33bc8ca0c99bff70aa174035096/PDF/2022/orga-nigramma2021_navigabile.pdf.

Ministero della Cultura. “Ufficio UNESCO.” Ultimo accesso 23 Gennaio 2024. <https://unesco.cultura.gov.it/>.

Moses, Olafenwa. “ImageAI: Object Detection.” Ultimo accesso 23 Gennaio 2024.
<https://github.com/OlafenwaMoses/ImageAI/blob/master/data-images/image3new.jpg>

MOTIONAL. “nuScenes”. Ultimo accesso 23 Gennaio 2024. <https://www.nuscenes.org/>.

- Mounier, Pierre. "Manifeste des Digital Humanities." *Journal des anthropologues*, n. 122-123 (2010). DOI: 10.4000/jda.3652.
- N.D. "Gb, la Camera dei Lord rinuncerà ai documenti scritti su pelle di vitello." *La Repubblica*, 11 Febbraio 2016.
https://www.repubblica.it/esteri/2016/02/11/foto/gb_la_camera_dei_lord_rinuncera_ai_documenti_scritti_su_pelle_di_vitello-133189845/1/.
- National History Museum of London. "The National History Museum of London." Ultimo accesso 23 Gennaio 2024. <https://data.nhm.ac.uk/>.
- National Information Standard Organization (NISO). *Understanding Metadata*. Bethesda, NISO Press, 2004.
<http://www.niso.org/publications/understanding-metadata>
- National Library of New Zealand. "Metadata Extractor." Ultimo accesso 23 Gennaio 2024.
<https://github.com/DIA-NZ/Metadata-Extraction-Tool>.
- Niola, Francesco. "L'Italia vuole deregolamentare l'IA generativa, ecco come." *Agenda Digitale*, 27 Novembre 2023. <https://www.agendadigitale.eu/sicurezza/privacy/litalia-vuole-deregolamentare-lia-generativa-ecco-come/>.
- Noakes, Drew. "Metadata-Extractor." Ultimo accesso 23 Gennaio 2024.
<https://github.com/drewnoakes/metadata-extractor>.
- Obermeyer, Ziad, Brian Powers, Christine Vogeli, and Sendhil Mullainathan. "Dissecting racial bias in an algorithm used to manage the health of populations," *Science* 366, n. 6464 (2019). DOI:10.1126/science.aax2342.
- Obeso, Abraham Montoya, Mireya S. García Vázquez, Alejandro A. Ramírez Acosta, and Jenny Benois-Pineau. "Connoisseur: classification of styles of Mexican architectural heritage with deep learning and visual attention prediction." *CBMI '17: Proceedings of the 15th International Workshop on Content-Based Multimedia Indexing*, (2017). DOI: 10.1145/3095713.3095730.
- Ocon, David. "Digitalising endangered cultural heritage in Southeast Asian cities: preserving or replacing?." *International Journal of Heritage Studies* 27, n. 10 (2021). DOI: 10.1080/13527258.2021.1883711.
- Open Preservation Foundation. "JHOVE." Ultimo accesso 23 Gennaio 2024.
<https://jhove.openpreservation.org/>.
- Open Preservation Foundation. "Open sustainable digital preservation." Ultimo accesso 23 Gennaio 2024.
<https://openpreservation.org/>.
- OpenCollective. "OpenSeadragon." Ultimo accesso 23 Gennaio 2024. <https://openseadragon.github.io/>.
- Ostertaga, Cecilia, and Marie Beurton-Aimar. "Matching ostraca fragments using a siamese neural network." *Pattern Recognition Letters*, (2020). DOI: 10.1016/j.patrec.2020.01.012.
- Parlamento Europeo. *Artificial intelligence in the context of cultural heritage and museums: Complex challenges and new opportunities* (2023).
[https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/747120/EPRS_BRI\(2023\)747120_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/747120/EPRS_BRI(2023)747120_EN.pdf).
- PASCAL VOC. "Dataset." Ultimo accesso 23 Gennaio 2024. <http://host.robots.ox.ac.uk/pascal/VOC/>.
- Pathak, Rachna, Anil Saini, Arnav Wadhwa, Himanshu Sharma, and Dhiraj Sangwan. "An object detection approach for detecting damages in heritage sites using 3-D point clouds and 2-D visual data." *Journal of Cultural Heritage* 48, (2021). DOI: 10.1016/j.culher.2021.01.002.

Penioni, Sofia, Simona Tiribelli, Marina Paolanti, Emanuele Frontoni, and Benedetta Giovanola. "Artificial Intelligence and Cultural Heritage: Design and Assessment of an Ethical Framework." *2023 IEEE International Symposium on Ethics in Engineering, Science and Technology (ETHICS)*, (2023). DOI: 10.5194/isprs-archives-XLVIII-M-2-2023-1149-2023.

Pigliapoco, Stefano, and Stefano Allegrezza. *La produzione e conservazione del documento digitale*. Macerata: EUM, 2008. ISBN: 978-88-6056-158-9.

Pontoppidan, Kristian. "GraphReader." Ultimo accesso 23 Gennaio 2024. <http://www.graphreader.com/>.

Porbital. "PlotDigitize." Ultimo accesso 23 Gennaio 2024. <https://plotdigitizer.com/>.

Project Mirador. "Mirador." Ultimo accesso 23 Gennaio 2024. <https://projectmirador.org>

Python. "Google-speech." Ultimo accesso 23 Gennaio 2024. <https://pypi.org/project/google-speech/>.

QaamGo. "Metadata2Go." Ultimo accesso 23 Gennaio 2024. <https://www.metadata2go.com/>.

Rane, Chinmayee, Mahadevan Seshasayee, Devi Sandeep Endluri, Wu Jian and Giles C. Lee. "ChartReader: Automatic Parsing of Bar-Plots." *2021 IEEE 22nd International Conference on Information Reuse and Integration for Data Science (IRI)*, (2021). DOI:10.1109/IRI51335.2021.00050.

Rane, Chinmayee. "ChartReader." Ultimo accesso 23 Gennaio 2024. <https://github.com/Cvrane/ChartReader..>

Răzvan, Gabriel Boboc, Elena Băutu, Florin Gîrbacia, Norina Popovici, and Dorin-Mircea Popovici. "Augmented Reality in Cultural Heritage: An Overview of the Last Decade of Applications." *Applied Sciences* 12, n. 19: 9859 (2022). DOI: 10.3390/app12199859.

Redmon, Joseph, Santosh Divvala, Ross Girshick, and Ali Farhadi. "You Only Look Once: Unified, Real-Time Object Detection." (2016). arxiv.org/pdf/1506.02640.

Reed, Jack. "Shimmy." Ultimo accesso 23 Gennaio 2024. <https://github.com/mejackreed/shimmy>.

Rehurek, Radim, and Sojka Petr. "Gensim-python framework for vector space modelling." *NLP Centre, Faculty of Informatics, Masaryk University, Brno (Czech Republic)* 3, n. 2 (2011). <https://radimrehurek.com/gensim/#>. Ultimo accesso Settembre 2023.

Ren, Shaoqing, Kaiming He, Ross Girshick, and Jian Sun. "Faster r-cnn: Towards realtime object detection with region proposal networks." (2015). <https://arxiv.org/abs/1506.01497>.

Rizk, Rodrigue, Dominick Rizk, Frederic Rizk, and Ashok Kumar. "A Hybrid Capsule Network-based Deep Learning Framework for Deciphering Ancient Scripts with Scarce Annotations: a Case Study on Phoenician Epigraphy." *2021 IEEE International Midwest Symposium on Circuits and Systems (MWSCAS)*, (2021). DOI: 10.1109/MWSCAS47672.2021.9531798.

Robinson, Glen. "Simple Annotation Server." Ultimo accesso 23 Gennaio 2024. <https://github.com/glenrobson/SimpleAnnotationServer>.

Rohatgi, Ankit. "WebPlotDigitizer." Ultimo accesso 23 Gennaio 2024. <https://automeris.io/WebPlotDigitizer/>.

Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-Net: Convolutional Networks for Biomedical Image Segmentation." (2015). <https://arxiv.org/abs/1505.04597>.

Roshka, Sebastian, and Vahid Mirjalili. *Python Machine Learning. Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow 2*. Birmingham: Pack Publishing. ed. Terza, 2019. ISBD: 978-1789955750.

- Rovella, Anna, Alexander Murzaku, Eugenio Cesario, Martin Critelli, Armando Bartucci, and Francesca M. C. Messiniti. "Analysis, evaluation and comparison of knowledge extraction tools in the Environmental and Health domain. A holistic approach." *Knowledge Organization and Management in the Domain of Environment and Earth Observation (KOMEEO) International Conference*, a cura di Antonietta Folino, and Roberto Guarasci. Ergon-Verlag, 2022. DOI:10.5771/9783956508752-121.
- Rovella, Anna, Roberto Guarasci, and Francesca Parisi. "Gestione e conservazione dei messaggi di posta elettronica." *AIDAinformazioni*, n. 1-2 (2020): 151-175.
<https://www.aidainformazioni.it/index.php/aidainformazioni/issue/view/4>.
- Rubini, Federico. "Il Forum di Davos, in diretta." *Corriere della Sera*, 18 Gennaio 2024,
https://www.corriere.it/economia/finanza/diretta-live/24_gennaio_18/forum-davos-diretta-herzog-prima-una-soluzione-politica-vogliamo-certezza-pace-d0ee8932-b5cf-11ee-8131-eee5ee6b3690.shtml?refresh_ce.
- Russakovsky, Olga, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg, and Li Fei-Fei. "ImageNet Large Scale Visual Recognition Challenge," *IJCV*, (2015). <https://doi.org/10.48550/arXiv.1409.0575>
- Safaa, Ismail S. I., Romany F. Mansour, Rasha M. Abd El-Aziz, and Ahmed I. Taloba. "Efficient E-Mail Spam Detection Strategy Using Genetic Decision Tree Processing with NLP Features." *Computational Intelligence and Neuroscience*, (2022). DOI: 10.1155/2022/7710005.
- Salarelli, Alberto. "International Image Interoperability Framework (IIIF): una panoramica." *JLIS* 8, n. 1 (2017). DOI: 10.4403/jlis.it-12090.
- Samvera Community. "Serverless-iiif." Ultimo accesso 23 Gennaio 2024.
<https://github.com/nulib/serverless-iiif>.
- Sapijaszko, Genevieve, and Wasfy B. Mikhael. "An Overview of Recent Convolutional Neural Network Algorithms for Image Recognition." *2018 IEEE 61st International Midwest Symposium on Circuits and Systems (MWSCAS)*, (2018). DOI: 10.1109/MWSCAS.2018.8623911.
- Sarcina, Franco. "L'incompiuta Decima Sinfonia di Beethoven portata a termine grazie all'Intelligenza Artificiale." *Il Sole 24 Ore*, 30 Settembre 2021. <https://www.ilsole24ore.com/art/l-incompiuta-decima-sinfonia-beethoven-portata-termine-grazie-all-intelligenza-artificiale-AEYmrXm>.
- Savva, Manolis, Nicholas Kong, Arti Chhajta, Li Fei-Fei, Maneesh Agrawala, and Jeffrey Heer. "ReVision: Automated Classification, Analysis and Redesign of Chart Images," *ACM User Interface Software & Technology (UIST. ACM)*, (2011). DOI:10.1145/2047196.2047247.
- Shinde, Pramila P., and Seema Shah. "A Review of Machine Learning and Deep Learning Applications." *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*, (2018). DOI: 10.1109/ICCUBEA.2018.8697857.
- Simonyan, Karen, and Andrew Zisserman. "Very Deep Convolutional Networks for Large-Scale Image Recognition." (2015). <https://arxiv.org/abs/1409.1556>.
- Siountri, Konstantina, and Christos-Nikolaos Anagnostopoulos. "The Classification of Cultural Heritage Buildings in Athens Using Deep Learning Techniques." *Heritage* 6, n. 4 (2023). DOI: 10.3390/heritage6040195.
- Slack Technologies. "Slack." Ultimo accesso 23 Gennaio 2024. <https://app.slack.com/>.
- Society of American Archivists. *Encoded Archival Description (EAD3)* (2019).
<https://www.loc.gov/ead/EAD3taglib/EAD3.html>.

- Soler Garrido, Josep, Delia Fano Yela, Cecilia Panigutti, Henrik Junklewitz, Ronan Hamon, Tatjana Evas, Antoine-Alexandre André, Salvatore Scalzo. *EUR 31518 EN Analysis of the preliminary AI standardization work plan in support of the AI Act* (2023). <https://op.europa.eu/en/publication-detail/-/publication/b14d9c86-faa3-11ed-a05c-01aa75ed71a1/language-en>.
- Spikerog. "Extract Metadata." Ultimo accesso 23 Gennaio 2024. <https://www.extractmetadata.com/>.
- Stanford University Human-Centered Artificial Intelligence. *Artificial Intelligence Index Report* (2023). https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI_AI-Index-Report_2023.pdf
- Swann, GM Peter. *The Economics of Standardization: An Update. Report for the UK Department of Business, Innovation and Skills (BIS)* (2010). https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/32444/10-1135-economics-of-standardization-update.pdf.
- Szabo, Liliana, Zahra Raisi-Estabragh, Ahmed Salih, Celeste McCracken, Esmeralda Ruiz Pujadas, Polyxeni Gkontra, Mate Kiss, Pal Maurovich-Horvath, and Hajnalka Vago. "Clinician's guide to trustworthy and responsible artificial intelligence in cardiovascular imaging," *Frontiers in Cardiovascular Medicine*, (2022). DOI: 10.3389/fcvm.2022.1016032.
- Szegedy, Christian, Vincent Vanhoucke, Sergey Ioffe, Jonathon Shlens, and Zbigniew Wojna. "Rethinking the Inception Architecture for Computer Vision." (2015). <https://arxiv.org/abs/1512.00567>.
- Szegedy, Christian, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. "Going deeper with convolutions." (2014). <https://arxiv.org/abs/1409.4842>.
- Tariq, Khizer. *Pakistani Cultural Clothes*. Ultimo accesso 23 Gennaio 2024. <https://www.kaggle.com/datasets/khizertariq/pakistani-cultural-clothes>.
- Tartaro, Alessio. "Towards European Standards supporting the AI Act: alignment challenges on the path to Trustworthy AI." *Proceedings of the AISB Convention 2023*, (2023), <https://aisb.org.uk/wp-content/uploads/2023/05/aisb2023.pdf>.
- Tesconi, Cristian. *Intelligenza Artificiale* (2023). ISBD: 979-8862365245.
- The Digital Innovation Group, Arizona State University. "cidoc-crm-neo4j." Ultimo accesso 27 Gennaio 2024. <https://github.com/diging/cidoc-crm-neo4j>.
- The National Archives of UK. "DROID." Ultimo accesso 23 Gennaio 2024. <https://www.nationalarchives.gov.uk/information-management/manage-information/preserving-digital-records/droid/>.
- Thoma, Martin. "PyPDF2". Ultimo accesso 23 Gennaio 2024. <https://github.com/py-pdf/pypdf/tree/3.0.0>.
- Thornton, Katherine, and Kenneth Seals-Nutt. "Science Stories: Using IIF and Wikidata to Create a Linked-Data Application." *International Workshop on the Semantic Web 2180*, (2018). <https://ceur-ws.org/Vol-2180/paper-68.pdf>.
- Tkaczyk, Dominika, Szostek Pawel, Fedoryszak Mateusz, Dendek Piot Jan, and Bolikowski Lukasz. "CERMINE: automatic extraction of structured metadata from scientific literature." *International Journal on Document Analysis and Recognition* 18, n. 4 (2015). DOI: 10.1007/s10032-015-0249-8.
- Trabelsi, Asma, Sébastien Warichet, Yassine Aajaoun, and Séverine Soussilane. "Evaluation of the efficiency of state-of-the-art Speech Recognition engines." *26th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems*, (2022). DOI:/10.1016/j.procs.2022.09.534.

- Uijlings, Jaspert, K. E. Van De Sande, T. Gevers, and A. W. Smeulders. "Selective search for object recognition." *International Journal of Computer Vision* 104, n. 2 (2013).
<https://ivi.fnwi.uva.nl/isis/publications/bibtexbrowser.php?key=UijlingsIJCV2013&bib=all.bib>
- United Nations for Education, Science and Culture (UNESCO) Glossary. "Cultural Heritage." Ultimo accesso 23 Gennaio 2024. https://uis.unesco.org/en/glossary/cultural_heritage.
- United Nations for Education, Science and Culture (UNESCO). "Dive into intangible cultural heritage." Ultimo accesso 23 Gennaio 2024. <https://ich.unesco.org/en/dive>.
- United Nations for Education, Science and Culture (UNESCO). "Intangible Heritage List." Ultimo accesso 23 Gennaio 2024. <https://ich.unesco.org/en/state/italy-IT?info=elements-on-the-lists>.
- United Nations for Education, Science and Culture (UNESCO). "UNESCO Digital Library." Ultimo accesso 23 Gennaio 2024. <https://unesdoc.unesco.org/>
- United Nations for Education, Science and Culture (UNESCO). "UNESCO in brief." Ultimo accesso 23 Gennaio 2024. <https://unesco.org/en/brief>.
- United Nations for Education, Science and Culture (UNESCO). "World Heritage List." Ultimo accesso 23 Gennaio 2024. https://whc.unesco.org/en/list/?search=&id_states=it&components=0&order=country.
- United Nations for Education, Science and Culture (UNESCO). *Convention Concerning the Protection of the World Cultural and Natural Heritage* (1972). <https://whc.unesco.org/en/conventiontext/>.
- United Nations for Education, Science and Culture (UNESCO). *Convention for the Safeguarding of the Intangible Cultural Heritage* (2003). <https://ich.unesco.org/en/convention>.
- United Nations for Education, Science and Culture (UNESCO). *Convenzione per la salvaguardia del patrimonio culturale immateriale*, cit., tradotto da Commissione Nazionale Italiana per UNESCO. <https://ich.unesco.org/doc/src/00009-IT-PDF.pdf>.
- United Nations for Education, Science and Culture (UNESCO). *Recommendation on the Safeguarding of Traditional Culture and Folklore* (1989). <https://www.unesco.org/en/legal-affairs/recommendation-safeguarding-traditional-culture-and-folklore>.
- United Nations for Education, Science and Culture (UNESCO). *UNESCO Universal Declaration on Cultural Diversity* (2001). <https://en.unesco.org/about-us/legal-affairs/unesco-universal-declaration-cultural-diversity>.
- United Nations for Education, Science and Culture Institute for Statistics (UNESCO UIS), and UNESCO Culture Sector. *2009 UNESCO Framework for Cultural Statistics (FCS)* (2009). https://uis.unesco.org/sites/default/files/documents/unesco-framework-for-cultural-statistics-2009-en_0.pdf. Questo documento costituisce una versione revisionata di quello edito nel 1986.
- United Nations for Education, Science and Culture Organization (UNESCO). *Recommendation on the Ethics of Artificial Intelligence* (2021). <https://unesdoc.unesco.org/ark:/48223/pf0000380455>.
- United Nations for Education, Science and Culture Organization (UNESCO). "World Heritage Online Map Platform." Ultimo accesso 23 Gennaio 2024.
<https://experience.arcgis.com/experience/aeae427b3c5045a6a4c72a93ab1d1280>
- United Nations Organization (ONU). *Interim Report: Governing AI for Humanity* (2023).
<https://www.un.org/ai-advisory-body>.
- University of Central Florida. *UFC Dataset* (2013). <https://www.crcv.ucf.edu/data/UCF101.php>.

- Vaishali, Kherdekar A., and Sachin A. Naik. "Speech Recognition System Approaches, Techniques And Tools For Mathematical Expressions: A Review." *International Journal of Scientific & Technology Research* 8, n. 8 (2019). <https://www.ijstr.org/final-print/aug2019/Speech-Recognition-System-Approaches-Techniques-And-Tools-For-Mathematical-Expressions-A-Review-.pdf>
- Verschoof-van der Vaart, Wouter, and Karsten Lambers. "Learning to Look at LiDAR: The Use of R-CNN in the Automated Detection of Archaeological Objects in LiDAR Data from the Netherlands." *Journal of Computer Applications in Archeology* 2, n. 1 (2019). DOI: 10.5334/jcaa.32.
- Vishal Dineshkumar Soni, "Speech Recognition: Transcription and transformation of human speech," *International Journal on Integrated Education* 2, n. 6 (2019). DOI: 10.31149/ijie.v2i6.497.
- W3C. "Metadata." Ultimo accesso 23 Gennaio 2024. <https://www.w3.org/DesignIssues/Metadata.html>.
- Wang, Changhan, Morgane Riviere, Ann Lee, Anne Wu, Chaitanya Talnikar, Daniel Haziza, Mary Williamson, Juan Pino, Emmanuel Dupoux. "VoxPopuli: A Large-Scale Multilingual Speech Corpus for Representation Learning, Semi-Supervised Learning and Interpretation." (2021). arXiv:2101.00390v2.
- Wang, Deng, Xiaodong Wang, and Shaohe Lv. "An Overview of End-to-End Automatic Speech Recognition." *Symmetry* 11, (2019). DOI:10.3390/sym11081018.
- Weiser, Benjamin. "Here's What Happens When Your Lawyer Uses ChatGPT." *New York Times*, 27 Maggio 2023, <https://www.nytimes.com/2023/05/27/nyregion/avianca-airline-lawsuit-chatgpt.html>.
- Wilkinson, Mark D. "The FAIR Guiding Principles for scientific data management and stewardship." *Scientific Data* 3, (2016). DOI: 10.1038/sdata.2016.18.
- WIT-AI. "PyWit." Ultimo accesso 23 Gennaio 2024. <https://github.com/wit-ai/pywit>.
- Xie, Saining, Chen Sun, Jonathan Huang, Zhuowen Tu, and Kevin Murphy. "Rethinking Spatiotemporal Feature Learning: Speed-Accuracy Trade-offs in Video Classification." (2018). <https://arxiv.org/abs/1712.04851>.
- Yinhan, Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. "RoBERTa : A Robustly Optimized BERT Pretraining Approach." (2019). arxiv.org/pdf/1907.11692.
- Zellers, Rowan, Yonatan Bisk, Ali Farhadi, and Yejin Choi. "From Recognition to Cognition: Visual Commonsense Reasoning." (2019). <https://arxiv.org/abs/1811.10830>.
- Zeng, Marcia Lei. "Semantic enrichment for enhancing LAM data and supporting digital humanities. Review article." *El profesional de la información* 28, n. 1(2019). DOI : 10.3145/epi.2019.ene.03.
- Zeng, Yuan, Yi Gong, and Xiangrui Zeng. "Controllable digital restoration of ancient paintings using convolutional neural network and nearest neighbor." *Pattern Recognition Letters* 133, (2020). DOI: 10.1016/j.patrec.2020.02.033.
- Zhang, Yiming, Ying Weng, and Jonathan Lund. "Applications of Explainable Artificial Intelligence in Diagnosis and Surgery." *Diagnostics* 12, n. 2 (2022). DOI: 10.3390/diagnostics12020237.
- Zhenzhong, Lan, Mingda Chen, Sebastian Goodman, Kevin Gimpel, Piyush Sharma, and Radu Soricut. "ALBERT: A Lite BERT for Self-supervised Learning of Language Representations." (2020). <https://arxiv.org/abs/1909.11942>.
- Zhong-Qiu, Zhao, Xu Shou-Tao, and Wu Xindong. "Object Detection with Deep Learning: A Review." (2019). <https://arxiv.org/abs/1807.05511>.

Zhou, Jun Yuhang Lu, Kang Zheng, Karen Smith, Colin Wilder, and Song Wang. "Design Identification of Curve Patterns on Cultural Heritage Objects: Combining Template Matching and CNN-based Re-Ranking." (2018). <https://arxiv.org/abs/1805.06862>.

Zoph, Barret, Vijay Vasudevan, Honathon Shlens, and Quoc V.Le. "Learning Transferable Architectures for Scalable Image Recognition." (2018). arXiv:1707.07012v4.

APPENDICE 1 – Scheda descrittiva in VRA Core de La Scuola di Atene

```
<?xml version="1.0" encoding="utf-8"?>
<vra xmlns="http://www.vraweb.org/vracore4.htm" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.vraweb.org/vracore4.htm http://www.loc.gov/standards/vracore/vra-strict.xsd">
  <work id="w_39" source="Core 4 Sample Database (VCat)" refid="39">
    <agentSet>
      <display>Raphael (Italian painter, 1483-1520)</display>
      <notes/>
      <agent>
        <name vocab="ULAN" refid="500023578" type="personal">Raphael</name>
        <dates type="life">
          <earliestDate>1483</earliestDate>
          <latestDate>1520</latestDate>
        </dates>
        <culture>Italian</culture>
        <role>painter</role>
      </agent>
    </agentSet>
    <dateSet>
      <display>1509-1510 (creation)</display>
      <date type="creation">
        <earliestDate>1509</earliestDate>
        <latestDate>1510</latestDate>
      </date>
    </dateSet>
    <locationSet>
      <display>Vatican, The (Rome, Holy See, Lazio, Italy)</display>
      <notes>Apostolic Palace, Stanza della segnatura</notes>
      <location type="repository">
        <name type="geographic" vocab="ULAN" refid="500311298" extent="site">Vatican, The</name>
        <name type="geographic" vocab="TGN" refid="7001168" extent="nation">Rome, Holy See</name>
        <name type="geographic" vocab="TGN" refid="7003080" extent="region">Lazio</name>
        <name type="geographic" vocab="TGN" refid="1000080" extent="nation">Italy</name>
        <name type="geographic" vocab="TGN" refid="1000003" extent="continent">Europe</name>
      </location>
    </locationSet>
    <titleSet>
      <display>School of Athens</display>
      <title type="cited" pref="true" xml:lang="en">School of Athens</title>
      <title type="cited" pref="false" xml:lang="en">Philosophy</title>
      <title type="cited" pref="false" xml:lang="it">Scuola di Atene</title>
    </titleSet>
    <worktypeSet>
      <display>paintings (visual works); frescoes (paintings)</display>
      <worktype vocab="AAT" refid="300177433">fresco (painting)</worktype>
    </worktypeSet>
  </work>
  <image id="i_141" href="http://www.core.vraweb.org/examples/html/example046_full.html" refid="141" source="VRA Core Oversight Committee,
">
  <relationSet>
```

```
<relation type="imageOf" refid="39" source="Core 4 Sample Database (VCat)"/>
</relationSet>
<titleSet>
  <display>Overall view in Stanze including doorway</display>
  <title type="generalView">Overall view in Stanze including doorway</title>
</titleSet>
<worktypeSet>
  <display>digital image</display>
  <notes/>
  <worktype/>
</worktypeSet>
</image>
</vra>
```

APPENDICE 2 – Esempio del file JSON-LD per il modello di *Collection* IIF

```
{
  "@context": "http://iiif.io/api/presentation/3/context.json",
  "id":
  "https://gist.githubusercontent.com/martincritelli/358a15ab518ef0c448ee429afb3850a1/raw/63ab832d69bbc0a7e6e909487dc30ff999848dbc/collection_sala_paolina_martin.json",
  "type": "Collection",
  "label": {
    "it": [
      "Sala 1 - Sala della Paolina"
    ]
  },
  "requiredStatement": {
    "label": {
      "en": [
        "Attribution"
      ]
    },
    "value": {
      "it": [
        "Galleria Borghese"
      ]
    }
  },
  "rights": "http://creativecommons.org/licenses/by/4.0/",
  "summary": {
    "it": [
      "La Sala I XVI secolo."
    ]
  },
  "provider": [
    {
      "id": "https://galleriaborghese.beniculturali.it/",
      "type": "Agent",
      "label": {
        "it": [
          "Galleria Borghese"
        ]
      },
      "homepage": [
        {
          "id": "https://galleriaborghese.beniculturali.it/",
          "type": "Text",
          "label": {
            "it": [
              "sito internet della Galleria Borghese"
            ]
          },
          "format": "text/html"
        }
      ]
    }
  ],
  "thumbnail": [
```



```

    {
      "id": "https://6012.cophilab-
cloud.ilc.cnr.it/share/tirocinio/sala_panoramica.png/full/full/0/default.jpg",
      "type": "Image",
      "format": "image/png",
      "service": [
        {
          "id": "https://6012.cophilab-cloud.ilc.cnr.it/share/tirocinio/sala_panoramica.png",
          "type": "ImageService2",
          "profile": "https://6012.cophilab-
cloud.ilc.cnr.it/share%2Ftirocinio%2Fsala_panoramica.png/level2.json"
        }
      ]
    }
  ],
  "items": [
    {
      "id":
      "https://gist.githubusercontent.com/martincritelli/6bd9c41e37db931193a596347e9a4b84/raw/46293673b051
b801cef3492e623d842fbcfc38ed/manifest_sala1_canvas_opere",
      "type": "Manifest",
      "label": {
        "it": [
          "Sala 1 in 360°"
        ]
      },
      "thumbnail": [
        {
          "id": "https://6012.cophilab-
cloud.ilc.cnr.it/share/tirocinio/sala_panoramica.png/full/full/0/default.jpg",
          "type": "Image",
          "format": "image/png",
          "service": [
            {
              "id": "https://6012.cophilab-cloud.ilc.cnr.it/share/tirocinio/sala_panoramica.png",
              "type": "ImageService2",
              "profile": "https://6012.cophilab-
cloud.ilc.cnr.it/share%2Ftirocinio%2Fsala_panoramica.png/level2.json"
            }
          ]
        }
      ]
    },
    {
      "id":
      "https://gist.githubusercontent.com/martincritelli/476e6f986aa4e43e18c4ea15f987dee2/raw/e669b294cd66d
68c6bafeb4ca14bfa423c67f763/manifest_erma_bacco.json",
      "type": "Manifest",
      "label": {
        "it": [
          "Erma di Bacco"
        ]
      },
      "thumbnail": [
        {

```

```

    "id": "https://6012.cophilab-cloud.ilc.cnr.it/share/tirocinio/Valadier-Erma-
Bacco.tif/full/full/0/default.jpg",
    "type": "Image",
    "format": "image/png",
    "service": [
      {
        "id": "https://6012.cophilab-cloud.ilc.cnr.it/share/tirocinio/Valadier-Erma-Bacco.tif",
        "type": "ImageService2",
        "profile": "http://6012.cophilab-cloud.ilc.cnr.it/share%2Ftirocinio%2FValadier-Erma-
Bacco.tif/level2.json"
      }
    ]
  }
]
}
]
}
]
}
]
}
]
}
}

```

APPENDICE 3 – Risultati video activity recognition

File	RetinaNet		S3D	
	Attività	Percentuale	Attività	Percentuale
Il liutaio di Cremona star della rete.mp4	making pizza	14%	bookbinding	33%
making_pizza.mp4	making pizza	79%	making pizza	99%
making_pizza_2.mp4	making pizza	8%	tasting beer	97%
muro_secco.mp4	abseiling	20%	digging	73%
processione_varia.mp4	archery	22%	diving cliff	47%
Pupi siciliani-Teatro dei pupi di Siracusa.mp4	smoking	12%	zumba	64%
pupi_1.mp4	motorcycling	38%	playing didgeridoo	28%
pupi_2.mp4	spray painting	14%	playing harp	97%
pupi_3.mp4	belly dancing	40%	playing cricket	38%
pupi_4.mp4	rock climbing	14%	climbing ladder	73%
Transumanza.mp4	riding mule	46%	riding or walking with horse	66%
Umberto Lanaro Liutaio.mp4	cooking sausages	6%	making jewelry	62%
Viaggio in Italia nel Patrimonio Immateriale UNESCO_ Alpinismo.mp4	bench pressing	29%	dancing ballet	47%
Viaggio in Italia nel Patrimonio Immateriale	dunking basketball	27%	dancing ballet	41%

UNESCO_ arte musicale dei suonatori di corno da caccia.mp4				
Viaggio in Italia nel Patrimonio Immateriale UNESCO_ il canto a tenore sardo.mp4	bench pressing	33%	dancing ballet	44%
Viaggio in Italia nel Patrimonio Immateriale UNESCO_ l'arte dei muretti a secco.mp4	air drumming	31%	dancing ballet	33%
Viaggio in Italia nel Patrimonio Immateriale UNESCO_ l'arte del pizzaiuolo napoletano.mp4	dunking basketball	31%	dancing ballet	53%
Viaggio in Italia nel Patrimonio Immateriale UNESCO_ l'arte della perla di vetro.mp4	air drumming	25%	dancing ballet	53%
Viaggio in Italia nel Patrimonio Immateriale UNESCO_ l'Opera dei Pupi Siciliani.mp4	dunking basketball	27%	dancing ballet	41%
Viaggio in Italia nel Patrimonio Immateriale UNESCO_ la dieta mediterranea.mp4	bench pressing	31%	dancing ballet	40%
Viaggio in Italia nel Patrimonio Immateriale UNESCO_ la perdonanza celestiniana.mp4	bench pressing	32%	dancing ballet	46%
Viaggio in Italia nel Patrimonio Immateriale	bench pressing	30%	dancing ballet	40%

UNESCO_ la pratica agricola della vite ad alberello.mp4				
Viaggio in Italia nel Patrimonio Immateriale UNESCO_ la transumanza.mp4	bench pressing	2%	dancing ballet	49%
Viaggio in Italia nel Patrimonio Immateriale UNESCO_ le feste delle grandi macchine a spalla.mp4	dunking basketball	31%	dancing ballet	50%
Viaggio in Italia nel Patrimonio Immateriale UNESCO_ saper fare liutario della tradizione cremonese (1).mp4	dunking basketball	28%	dancing ballet	45%
Viaggio in Italia nel Patrimonio Immateriale UNESCO_ saper fare liutario della tradizione cremonese.mp4	dunking basketball	28%	dancing ballet	45%
Viaggio in Italia nel Patrimonio Unesco_ arte rupestre della Val Camonica.mp4	bench pressing	31%	dancing ballet	26%
Viaggio in Italia nel Patrimonio Unesco_ colline del Prosecco di Conegliano e Valdobbiadene.mp4	bench pressing	29%	dancing ballet	31%
Viaggio in Italia nel Patrimonio Unesco_ i sassi	air drumming	30%	dancing ballet	27%

e il parco delle chiese rupestri di Matera.mp4				
Viaggio in Italia nel Patrimonio Unesco_ la città di Vicenza e le ville del Palladio in Veneto.mp4	bench pressing	30%	dancing ballet	25%
Viaggio in Italia nel Patrimonio Unesco_ Trulli di Alberobello.mp4	air drumming	30%	dancing ballet	29%
Viaggio in Italia nel Patrimonio Unesco_ Val d'Orcia.mp4	bench pressing	28%	dancing ballet	28%
“Lo scrigno della memoria”.mp4	tasting beer	21%	giving or receiving award	22%

APPENDICE 4 – Estrazione automatica delle informazioni dai testi

Grazie al progressivo sviluppo di tecnologie sempre più affidabili e accessibili, sono stati implementati strumenti in grado di effettuare con maggiore accuratezza le attività descritte precedentemente. Il vasto panorama tecnologico attualmente a disposizione impone di limitare la trattazione, nel presente capitolo, ad un numero circoscritto di strumenti di estrazione automatica delle informazioni, prediligendo soluzioni di tipo open source, tra le quali saranno selezionati i tool più idonei per le attività di sperimentazione del lavoro di tesi. Per una più approfondita enucleazione, sarà seguito il criterio imposto per la descrizione delle tecniche di estrazione automatica, suddividendo gli strumenti per tipologia documentale di applicazione e al tipo di attività processabile:

- Testi:
 - Estrazione di entità testuali
 - Estrazione di tabelle e immagini
 - Metadati embedded ed external
- Immagini:
 - Classi di oggetti e individui
 - Metadati embedded
- Audiovideo:
 - Trascrizione del parlato
 - Metadati embedded
 - Classi di oggetti e individui
 - Riconoscimento delle attività

L'estrazione automatica delle informazioni da testi concerne l'applicazione di tecniche in grado di indagare il testo sotto il profilo strutturale e semantico.

Si parla di *Natural Language Processing* (NLP) – in italiano generalmente espresso come *Trattamento Automatico del Linguaggio* (TAL) – per indicare la disciplina «*volta a sviluppare sistemi che eseguono dei compiti linguistici, simulando le competenze e i comportamenti linguistici degli esseri umani*»³¹⁹.

Il NLP si compone di una serie di tecniche per l'elaborazione di grandi quantità di dati volte a espletare attività afferenti al linguaggio umano come ad esempio, la traduzione automatica, l'estrazione di termini o la classificazione dei testi in base al contenuto. Oggetto di analisi sono,

³¹⁹ Elisabetta Ježek, and Rachele Sprugnoli. *Linguistica computazionale. Introduzione all'analisi automatica dei testi*, (Il Mulino, 2023), p. 22. ISBN: 978-8815290359.

dunque, dati testuali, non strutturati e semi-strutturati, in linguaggio naturale che possono essere analizzati per scopi e modalità differenti. In particolare, rientrano nella sfera applicativa di NLP:

- Analisi del sentimento (*Sentiment Analysis*). L'obiettivo è classificare le emozioni a partire dal contenuto di un testo. È largamente impiegata nell'industria cinematografica per analizzare le recensioni di film o, più genericamente, per analizzare i feedback degli utenti sui prodotti in commercio. Oltre che per le recensioni, la *sentiment analysis* è applicata per valutare le emozioni delle persone a partire dal linguaggio impiegato distinguendo tra sentimenti positivi, negativi o neutrali.
- La creazione di assistenti virtuali. Si tratta di sistemi creati con la finalità di interagire con l'uomo attraverso il linguaggio naturale. Per essere in grado di comprendere e impiegare il linguaggio naturale, richiedono un addestramento su enormi quantità di tali dati testuali.
- Classificazione dei testi. Questa operazione include diverse attività di NLP volte a indagare il contenuto dei testi sotto molteplici aspetti (semantico, sintattico e morfologico etc.) per la classificarli in base alle tematiche di maggior rilievo rintracciabili all'interno.
- Traduzione automatica. Una delle primissime, e alquanto generali, applicazioni di NLP finalizzata alla traduzione automatica dei testi. È un'attività in progressivo miglioramento che richiede un'elevata conoscenza della struttura morfo-sintattica delle lingue per elaborare traduzioni accurate che tengano, anche, conto del contesto della frase originale.
- Costruzione di ontologie (*Ontology learning*)³²⁰. Con il termine *Ontology learning*³²¹ si suole indicare l'applicazione di tecniche di NLP, come ad esempio l'estrazione di termini e di relazioni, per la costruzione automatica di ontologie a partire dai testi in linguaggio naturale. I task di NLP per la costruzione automatica di ontologie possono essere organizzati all'interno di ciò che è meglio conosciuto con il termine *layer cake*³²². Riprendendo il concetto di *semantic web layer cake* dalle tecnologie semantiche del W3C³²³, il termine *layer cake* indica una classificazione per complessità dei sub task dediti alla creazione automatica di ontologie (Figura 37).

³²⁰ Nathalie Aussenac-Gilles, Brigitte Biebow, and Sylvie Szulman, "Revisiting ontology design: A methodology based on corpus analysis," *Proceedings of the International Conference on Knowledge Engineering and Knowledge Management (EKAW) 1937*, (2000), pp. 172–188. DOI: 10.1007/3-540-39967-4_13.

³²¹ Philipp Cimiano, Alexander Mädche, Steffen Staab, and Johanna Völker, "Ontology learning," in *Handbook on Ontologies*, ed. Staab Steffen, and Rudi Studer, (Berlin: Springer, 2004) 245-267.

³²² Philipp Cimiano, *Ontology Learning and Population from Text: Algorithms, Evaluation and Application*, (New York: Springer-Verlag, 2006).

³²³ DBpedia, "Semantic Web Stack," ultimo accesso 23 Gennaio 2024, https://dbpedia.org/page/Semantic_Web_Stack.

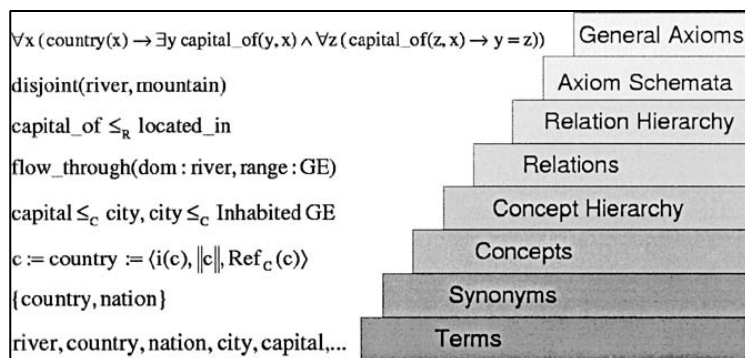


Figura 37: *Ontology learning layer cake*³²⁴.

Per le potenzialità espresse da tali tecniche, il NLP è strettamente connesso con l'Information Retrieval (IR)³²⁵ per la costruzione di sistemi in grado di soddisfare il bisogno informativo degli utenti mediante *query*³²⁶. In tale dominio, l'applicazione di NLP permette infatti di mitigare la frequente ambiguità e imprecisione nella morfologia di documenti e query, riconducendo varianti della stessa forma, sintattica, individuando le dipendenze che occorrono tra più termini e semantica ricorrendo a thesauri e vocabolari controllati.

Genericamente, un'attività di trattamento automatico del linguaggio si compone di un **input**, ovvero i dati da processare, il **modello computazionale**, l'oggetto di calcolo implementato per l'espletamento dell'attività e l'**output**, vale a dire il risultato elaborato sulla base del modello. In base alla specifica attività, gli algoritmi implementati all'interno dei modelli computazionali possono basarsi sull'apprendimento supervisionato o non supervisionato.

Nell'analisi dei testi, l'impiego di tecniche di NLP permette di raggiungere notevoli risultati in molteplici ambiti quali la classificazione automatica delle email³²⁷ e in particolare nell'individuazione di spam³²⁸, estrazione degli argomenti da testi di dimensioni anche estese (*topic modelling*), disambiguazione delle parole (*word sense disambiguation*) e così via.

Per gli scopi del presente elaborato, si procederà a descrivere alcuni tra i principali task di NLP, in particolare quelli *supervised learning-based* impiegati all'interno della parte applicativa.

³²⁴ Philipp Cimiano, *Ontology Learning and Population from Text: Algorithms, Evaluation and Application cit.*, p. 23.

³²⁵ Christopher D. Manning, Prabhakar Raghavan, and Hinrich Schütze, *Introduction to Information Retrieval* (Cambridge University Press. Online edition, 2009), p. 262. <https://nlp.stanford.edu/IR-book/pdf/irbookonlinereading.pdf>

³²⁶ «Interrogazione che l'utente di un sistema informatico o di una rete telematica fa, con i codici opportuni, al sistema o a un *data base*, alla ricerca di una specifica informazione.» Cfr. Enciclopedia Treccani, "Query," ultimo accesso 23 Gennaio 2024, www.treccani.it/vocabolario/query/.

³²⁷ Anna Rovella, Roberto Guarasci, and Francesca Parisi, "Gestione e conservazione dei messaggi di posta elettronica." *AIDAinformazioni*, n. 1-2 (2020):151-175. <https://www.aidainformazioni.it/index.php/aidainformazioni/issue/view/4>.

³²⁸ Safaa S. I. Ismail et al., "Efficient E-Mail Spam Detection Strategy Using Genetic Decision Tree Processing with NLP Features." *Computational Intelligence and Neuroscience*, (2022). Doi: 10.1155/2022/7710005.

Preprocessing

L'attività di *preprocessing* è propedeutica a qualunque tipologia di analisi del testo. In questa fase l'obiettivo è quello di rendere il testo originale processabile e quindi convertirlo in «sequenze di unità linguisticamente significative»³²⁹. Il task di *preprocessing* comprende diverse elaborazioni quali:

- Tokenizzazione
- Lower casing
- Rimozione delle stopwords
- Stemming
- Lemmatizzazione

La tokenizzazione l'operazione di partenza che consente di segmentare il testo e di estrarne i token, ossia parole semplici, contenuti. Similarmente, è possibile segmentare il testo in frasi (*sentence segmentation*). La riduzione in minuscolo dei caratteri (*lower casing*) è un processo attuato per semplificare analisi che richiedono uniformità dal punto di vista del segno e la cui distinzione tra caratteri maiuscoli e minuscoli comporterebbe la perdita di dati. In fase di pre-processamento, uno dei problemi con cui ci si confronta è il rumore informativo, spesso causato dalla presenza di parole prive di significato come articoli o preposizioni. In tale circostanza è necessario procedere alla rimozione delle cosiddette *stopwords*, ossia una lista di parole “vuote” non indispensabili ai fini dell'analisi. Questa lista varia in base alla lingua di riferimento e può contenere, a seconda del caso, diversi elementi: articoli, preposizioni, pronomi, congiunzione etc. Lo *stemming* è l'operazione mediante cui si troncano le parole rimuovendone gli affissi in modo da ricondurre le varianti morfologiche al tema di appartenenza. Ad esempio, avendo nel testo forme coniugate del verbo *scrivere* (*scrivevo, scrissi, scrivono*), queste verranno “stemmatizzate” per ottenere in tutti i casi il tema, ossia *scriv*. Infine, la *lemmatizzazione* permette di raggruppare le forme flesse di una parola sotto la propria unità di base, ovvero il lemma. Si tratta di un'operazione complessa poiché si trova a dover risolvere il problema legato all'ambiguità dettata dal contesto al fine di associare ciascuna parola al proprio lemma. Tornando all'esempio visto in precedenza, il lemma a cui saranno ricondotte le varie declinazioni sarà il verbo *scrivere*.

Il risultato di questa serie di operazioni è dunque un testo segmentato e disambiguato che può essere processato in base agli obiettivi prefissati.

³²⁹ Elisabetta Ježek, and Rachele Sprugnoli, *Linguistica computazionale. Introduzione all'analisi automatica dei testi cit.* p. 167.

Named Entity Recognition

Named Entity Recognition (NER) è l'attività volta a individuare le entità presenti nel testo e a classificarle in base alla categoria di appartenenza.

Le categorie alle quali i termini possono essere ricondotti, possono essere generiche oppure di dominio. Nella prima rientrano entità comuni come lo spazio, il tempo, nomi, (organizzazioni, persone etc.), mentre la categoria di dominio è specifica dell'ambito in cui il task viene applicato.

Principalmente si distinguono tre diversi approcci³³⁰ per l'elaborazione del NER:

- 1. Rule-based.** La corrispondenza con la categoria di appartenenza viene stabilita mediante la costruzione di set di regole e di modelli di classificazione. Questo approccio implica una conoscenza a priori del dominio e della lingua per stabilire regole efficaci oltre che la costruzione a priori di modelli coerenti.
- 2. Machine Learning.** Vengono applicati algoritmi di machine learning basati su modelli statistici per individuare e classificare raggruppamenti e relazioni nel testo. Si possono ulteriormente distinguere in questa categoria modelli di apprendimento supervisionato e non supervisionato.
- 3. Ibrido.** È l'approccio che impiega congiuntamente tecniche basate sulla costruzione di regole e algoritmi di ML.

Topic modelling

Il *Topic modelling* è l'attività che consente di identificare i contenuti di un testo o di un discorso. Uno dei modelli³³¹ impiegati per l'esecuzione di tale attività è il *Latent Dirichlet Allocation* (LDA), un modello probabilistico generativo in grado di analizzare grandi collezioni di documenti testuali per individuarne i topic. Il punto di partenza è l'assunzione, da parte del modello, che all'interno di una collezione di testi siano presenti più temi a cui una, o più, parole sono associate. Per una corretta elaborazione del task è necessario che i testi siano tokenizzati, normalizzati e le parole irrilevanti rimosse (*stopwords removal*). A queste operazioni di preparazione si aggiunge anche la creazione di bigrammi e trigrammi, ossia di coppie e terne di parole con una co-occorrenza maggiore all'interno del testo che possono contribuire a modellarne il significato. In LDA, la distribuzione di Dirichlet è applicata per definire la distribuzione dei topic all'interno di ogni documento e la probabilità che una parola sia associata ad un particolare tema.

³³⁰ Archana Goyal, Vishal Gupta, and Manish Kumar. "Recent Named Entity Recognition and Classification techniques: A systematic review," *Computer Science Review* 29, (2018), p. 5. DOI: 10.1016/j.cosrev.2018.06.001.

³³¹ Per un approfondimento circa i modelli adoperati in tale contesto si suggeriscono letture specifiche come, ad esempio, Pooja Kherwa, , and Poonam Bansal, "Topic Modeling: A Comprehensive Review," *EAI*, (2019). DOI: 10.4108/eai.13-7-2018.159623.

Le tecniche di riconoscimento automatico degli argomenti possono avere un ruolo fondamentale nella classificazione dei testi e costituiscono il nucleo in task come l'analisi dei contenuti dei social media o degli articoli di giornale o anche la personalizzazione dei contenuti proposti in rete. In ambito documentale, tali tecniche possono anche essere applicate nell'identificazione di dati personali e sensibili³³² da oscurare o comunque da rendere non accessibili a coloro che non sono autorizzati.

Strumenti di estrazione automatica delle informazioni applicati ai testi

All'interno di questa categoria rientrano strumenti in grado di processare ed estrarre le informazioni relative al contesto e al contenuto dei documenti testuali. Specificatamente per la prima categoria, l'obiettivo è quello di estrarre set di metadati descrittivi, tecnici e amministrativi che possono far emergere informazioni in grado di agevolare l'implementazione di servizi di conservazione e tutela dei beni documentali. Per questa attività allo stato attuale è possibile impiegare strumenti *web-based* come Metadata2Go³³³ o ExtractMetadata³³⁴, applicativi software tra cui Libextractor³³⁵ e Metadata-Extractor³³⁶ oppure librerie in python quale PyPDF2³³⁷.

Per ciò che concerne l'estrazione dei metadati dal contenuto delle risorse, esiste un numero ridotto di strumenti che sono sviluppati, in particolare, con l'obiettivo di processare articoli scientifici quali, ad esempio Grobid³³⁸ e Metatagger³³⁹.

Per le attività legate in questo progetto di dottorato sono stati applicati due specifici e ben distinti framework in grado di elaborare le due forme di estrazione di metadati: FITS (File Information Tool Set) per l'estrazione di metadati associati alla risorsa e CERMINE per l'estrazione di metadati dal contenuto.

Quest'ultimo, in particolare, in aggiunta all'estrazione elabora la conversione dei testi da pdf a xml che agevola i successivi processi di analisi tramite le tecniche di NLP.

FITS

Il FITS è stato sviluppato dalla Harvard Library nel 2009 in JAVA per l'estrazione di metadati *embedded* da documenti testuali, audiovisivi e immagini, l'identificazione del formato del file e la verifica di conformità del documento secondo le specifiche del formato di riferimento. È disponibile

³³² Tim Hutchinson, "Protecting privacy in the archives: Preliminary explorations of topic modeling for born-digital collections," *2017 IEEE International Conference on Big Data (Big Data)* 2251-55, (2018). Doi:10.1109/BigData.2017.8258177

³³³ QaamGo, "Metadata2Go," ultimo accesso 23 Gennaio 2024, <https://www.metadata2go.com/>

³³⁴ Spikerog, "Extract Metadata," ultimo accesso 23 Gennaio 2024, <https://www.extractmetadata.com/>.

³³⁵ GNU, "Libextractor," ultimo accesso 23 Gennaio 2024, <https://www.gnu.org/software/libextractor/>.

³³⁶ Drew Noakes, "Metadata-Extractor," ultimo accesso 23 Gennaio 2024, <https://github.com/drewnoakes/metadata-extractor>.

³³⁷ Martin Thoma, "PyPDF2," ultimo accesso 23 Gennaio 2024, <https://github.com/py-pdf/pypdf/tree/3.0.0>.

³³⁸ Patrice Lopez, "Grobid," ultimo accesso 23 Gennaio 2024, <https://github.com/kermitt2/grobid>.

³³⁹ Andrew McCallum, "Metatagger," ultimo accesso 23 Gennaio 2024, <https://github.com/iesl/rexa1-metatagger>.

sia in versione *stand-alone*, scaricando il file dell'applicativo, che come *web service* tramite Tomcat o JBoss. Il framework è strutturato come un "involucro" comprendente al proprio interno dodici tool, ognuno dei quali sviluppato esternamente e designato allo svolgimento di una specifica attività. Questi tool sono connessi all'interno del FITS tramite un *wrapper* che riceve in input il risultato derivante dal singolo tool e lo converte verso il formato FITS XML. Successivamente, attraverso un processo di normalizzazione, i singoli risultati convertiti vengono raggruppati insieme per formare il risultato finale del framework. L'output del FITS è un documento in formato FITS XML o XML standard, che può essere specificato in stringa di comando, organizzato in sezioni:

- 1) **Identification**: in questa sezione vengono eseguiti i tool che sono addetti all'identificazione del formato del file e del MIME type (indicazioni al browser su come gestire le estensioni dei file e su come aprire files con determinate estensioni). Insieme al metadato «identity», questa sezione riporta anche il nome dei tool usati per il riconoscimento il formato del file e la loro versione;
- 2) **Fileinfo**: è la sezione riguardante i metadati tecnici comuni a più formati come ad esempio «created» o «lastmodified». Ogni metadato riporta, insieme al valore di appartenenza, il tool che lo ha estratto insieme alla versione d'utilizzo. Generalmente, i tool eseguiti in questa sezione sono: JHOVE, TIKA, EXIFTOOL e OIS File Information;
- 3) **Filestatus**: è la sezione che contiene le informazioni sulla validità e formazione del documento «well-formed» e «valid». Questi metadati contengono un valore Booleano che può essere «true» o «false». JHOVE è il tool che viene eseguito in questa sezione;
- 4) **Metadata**: contiene i metadati tecnici suddivisi in: Audio elements, Document elements, Image elements, Text elements e Video elements. In questa sezione verranno eseguiti quei tool che elaborano l'estrazione di metadati e che supportano il formato del documento analizzato. Nel caso in cui il formato sia un video, i tool eseguibili sono: NLNZ Metadata Extractor Tool, EXIFTOOL, JHOVE e MediaInfo;
- 5) **ToolOutput**: contiene le informazioni sui tool attivati nel processo e la loro versione di sviluppo.
- 6) **Statistics**: contiene le informazioni sul tempo impiegato (in secondi) da ciascun tool, insieme alla relativa versione, per essere eseguito. Nei casi in cui invece un tool non dovesse essere eseguito, il valore riportato sarà «did not run».

Lo schema riportato nella Figura 38 offre una sintesi della struttura del FITS e dei tool incorporati che, nello specifico, sono:

- ADL Tool; sviluppato dalla stessa Harvard Library permette di riconoscere documenti audio;

- JHOVE³⁴⁰: consente l'identificazione, validazione ed estrazione di metadati tecnici da formati quali: jpg, wave, html, pdf, utf-8 etc.;
- DROID³⁴¹: identifica il formato del file tra oltre 1000 formati;
- EXIFTOOL³⁴²: identifica il file ed estrae metadati tecnici da molteplici formati tra cui: jpg, mp3, doc, pdf, hmtl etc.;
- NZNL Metadata Extractor Tool³⁴³: identifica il formato del file ed estrae metadati tecnici supportando una pluralità di formati come: jpg, wav, html,xml, pdf, doc etc.;
- FFIDENT³⁴⁴: identifica il formato del file tra il riconoscimento di formati quali ad esempio: xml, html, pdf, bmp, gif, avi, mp3 etc.;
- MediaInfo³⁴⁵: opera su documenti video identificandone il formato ed estraendone metadati tecnici;
- File Utility: è applicabile esclusivamente su Windows e permette d'identificare il formato del file tra oltre 1000 formati;
- OIS Audio Information: supporta i formati audio;
- OIS File Information: estrae metadati tecnici quali il nome, la grandezza o il percorso del file supportando qualsiasi tipo di formato;
- OIS XML Information: identifica il file in formato xml e ne estrae i metadati tecnici;
- TIKA³⁴⁶: identifica il formato del file tra oltre 1400 formati.

³⁴⁰ Open Preservation Foundation, "JHOVE," ultimo accesso 23 Gennaio 2024, <https://jhove.openpreservation.org/>.

³⁴¹ The National Archives of UK, "DROID," ultimo accesso 23 Gennaio 2024, <https://www.nationalarchives.gov.uk/information-management/manage-information/preserving-digital-records/droid/>.

³⁴² Phil Harvey, "ExifTool," ultimo accesso 23 Gennaio 2024, <https://exiftool.org/>.

³⁴³ National Library of New Zealand, "Metadata Extractor," ultimo accesso 23 Gennaio 2024, <https://github.com/DIA-NZ/Metadata-Extraction-Tool>.

³⁴⁴ Gary McGath, "Ffident," ultimo accesso 23 Gennaio 2024, <https://github.com/gmcgath/ffident>.

³⁴⁵ Java Native Access, "MediaInfo," ultimo accesso 23 Gennaio 2024, <https://github.com/java-native-access/jna>.

³⁴⁶ Apache TIKA, "Formats," ultimo accesso 23 Gennaio 2024, <https://tika.apache.org/1.20/formats.html>.

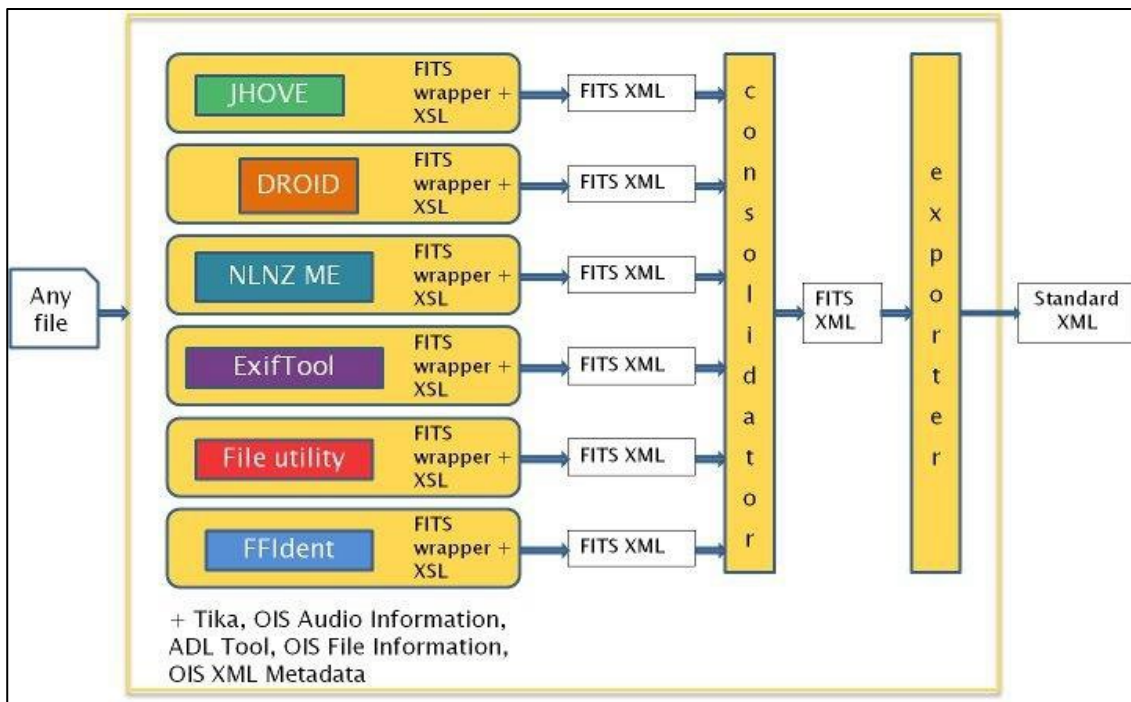


Figura 38: Struttura di FITS³⁴⁷.

Oltre a questi strumenti, è possibile incorporarne altri tramite la funzione sviluppatore, creando per ciascuna aggiunta un apposito *wrapper* in grado di convertire l'output secondo le specifiche del formato FITS XML³⁴⁸.

Riguardo al processo di elaborazione il framework è costruito per lavorare secondo quanto segue. Al momento di avvio, in base al formato del documento in input, ciascun tool potrà essere attivato singolarmente. Nel caso specifico di testi, con un documento in formato PDF strumenti quali ADL Tool o OIS Audio Information, designati per lavorare specificatamente su documenti sonori, non riconoscendo un formato adeguato non saranno attivati. Questa architettura consente di ottimizzare la fase di processamento e restituire un risultato nel minor tempo possibile. In fase di output, in base alla scelta operata dall'utente, il risultato sarà un documento in formato XML suddiviso nelle sezioni precedentemente descritte e tramite le quali sarà possibile evincere i tool azionati o meno per il processamento del documento in input.

Grazie al vasto numero di strumenti incorporati al proprio interno, con la conseguente possibilità di essere applicato su diverse tipologie documentali, il FITS è stato selezionato con l'obiettivo di estrarre i metadati associati alle risorse e ai relativi formati, operazione indispensabile per accertare che un documento sia integro e non abbia subito corruzioni di alcun genere.

³⁴⁷ Harvard Library, "File Information Tool Set (FITS)," ultimo accesso 23 Gennaio 2024, <https://projects.iq.harvard.edu/fits/fits-processing>.

³⁴⁸ Per un approfondimento tecnico si rimanda alla guida ufficiale. Cfr. Harvard Library, "Adding tool," ultimo accesso 23 Gennaio 2024, <https://projects.iq.harvard.edu/fits/fits/adding-tool>.

CERMINE

Content ExtRactor and MINEr, CERMINE, è un framework open source sviluppato in Java per l'estrazione automatica di metadati e di contenuti da documenti scientifici nativi digitali e, in particolare, afferenti al dominio medico. Tuttavia, il framework consente l'addestramento personalizzato per l'estrazione di metadati da differenti domini. CERMINE è sviluppato per funzionare in moduli separati tra loro distinti in: *Basic structure extraction*, *Metadata extraction* e *Bibliography extraction*. Questi moduli sono caratterizzati da diversi algoritmi di machine learning per l'elaborazione di specifiche attività. La Figura 39 fornisce uno schema riassuntivo dei moduli con annessi processi e algoritmi applicati.

Path	Step	Goal	Implementation
A. Basic structure extraction	A1. Character extraction	Extracting individual characters along with their page coordinates and dimensions from the input PDF file	iText library
	A2. Page segmentation	Constructing the document's geometric hierarchical structure containing (from the top level) pages, zones, lines, words and characters, along with their page coordinates and dimensions	Enhanced Docstrum
	A3. Reading order resolving	Determining the reading order for all structure elements	Bottom-up heuristic-based
	A4. Initial zone classification	Classifying the document's zones into four main categories: <i>metadata</i> , <i>body</i> , <i>references</i> and <i>other</i>	SVM
B. Metadata extraction	B1. Metadata zone classification	Classifying the document's zones into specific metadata classes	SVM
	B2. Metadata extraction	Extracting atomic metadata information from labelled zones	Simple rule-based
C. Bibliography extraction	C1. Reference strings extraction	Dividing the content of <i>references</i> zones into individual reference strings	K-means clustering
	C2. Reference parsing	Extracting metadata information from references strings	CRF

Figura 39: Composizione del workflow di CERMINE³⁴⁹.

L'attività preliminare del workflow di CERMINE, propedeutica alla riuscita dell'intero flusso, è costituita dal modulo *Basic structure extraction*, all'interno del quale prende avvio la segmentazione geometrica della struttura del documento, indispensabile per il riconoscimento del contenuto informativo. Questa operazione, denominata *Character extraction* processa il documento individuandone le coordinate geometriche e le dimensioni da cui partire per l'elaborazione del *Page segmentation* e *Reading order resolving* per l'individuazione della struttura gerarchica, le pagine, paragrafi, righe e caratteri, contestualmente alla definizione del corretto ordine di lettura del testo. Infine, l'ultimo processo all'interno di questo modulo iniziale è costituito dall'*Initial zone classifier* che ha come obiettivo la divisione del testo e la classificazione in zone distinte in: *metadata*, *references*, *body* e *other*. Il risultato di questa operazione costituisce la base di partenza per l'elaborazione dei processi dei moduli successivi: *Metadata extraction* e *Bibliography extraction*. Questi moduli

³⁴⁹ Dominika Tkaczyk et al., "CERMINE: automatic extraction of structured metadata from scientific literature." *cit.* p.321.

processano esclusivamente le zone di interesse, rispettivamente *metadata* e *references*, per riconoscere ed estrarre metadati descrittivi e bibliografici. Nel modulo *Metadata extraction* il processo iniziale, *Metadata zone classification*, ha come obiettivo l'analisi delle zone precedentemente classificate come *metadata* per l'assegnazione di classi di metadati quali, ad esempio, *title*, *author*, *date* etc. Una volta individuati questi elementi, il processo di *Metadata extraction* provvede ad estrarli nel formato XML JATS, specifico per la codifica dei metadati della letteratura scientifica. Procedimento analogo accade nell'ultimo modulo dove, a partire dalle zone etichettate come *references*, il processo di *Reference strings extraction* elabora l'individuazione, estrazione e concatenazione delle singole stringhe di citazione mediante un algoritmo di clustering. Infine, il processo di *Reference string parser* compie l'estrazione dei metadati bibliografici nel formato BibTex. Questa attività prevede la iniziale suddivisione delle stringhe in singoli token e l'impiego di un algoritmo di *Conditional Random Fields* (CRF) per la classificazione dei token in classi di metadati come *first_name*, *email*, *surname* etc.

In aggiunta a quanto descritto sinora, CERMINE esegue anche l'estrazione delle immagini dal testo tra le quali, talvolta, è possibile annoverare anche tabelle e grafici, e la conversione dell'intero testo in formato XML.

Trattamento automatico del linguaggio

L'applicazione dei due strumenti appena descritti è indispensabile per convertire i documenti in un formato testuale, come l'XML o il TXT, che possa essere trattato con tecniche di *Natural Language Processing* (NLP) per l'analisi e l'estrazione delle entità testuali. Infatti, per l'applicazione di queste tecniche è essenziale che il testo sia in un formato *machine-readable* come l'XML o, in ogni caso, che non contenga limiti imposti dalle codifiche di formati come il PDF o il DOCX.

Per il trattamento automatico del linguaggio esistono allo stato attuale diverse librerie Python in grado di elaborare le innumerevoli operazioni sui testi.

spaCy. Una libreria open source per l'elaborazione delle principali operazioni di NLP come la tokenizzazione, il tagging del Part-of-Speech (POS), lemmatizzazione, Named Entity Recognition (NER) etc., su grandi quantità di testi. Si tratta di una libreria tra le più usate e continuamente aggiornate che, allo stato attuale, supporta oltre 73 lingue con la possibilità di analizzare i testi multilingua. Il documento in input deve essere un testo in formato TXT e, ai fini di una analisi quanto più precisa, occorre specificare il modello linguistico scegliendo tra quelli pre-addestrati presenti in spaCy o importandone uno proprio. Per l'elaborazione dei diversi task, la libreria presenta una pipeline semplice attraverso la quale all'utente è richiesto solo di caricare il modello di base e specificare l'azione da compiere³⁵⁰.

³⁵⁰ Per completezza sulle operazioni possibili si rimanda alla pagina web di riferimento. Cfr. EXPLOSION. "linguistic-features," ultimo accesso 23 Gennaio 2024, <https://spacy.io/usage/linguistic-features>.

NLTK. Natural Language Toolkit (NLTK) è una libreria open source per l'analisi dei dati linguistici attraverso l'accesso ad oltre 50 corpora e risorse lessicali e alle differenti operazioni di trattamento del testo: classificazione, tokenizzazione tagging etc. Simili per alcune funzionalità, spaCy e NLTK si differenziano sotto l'aspetto metodologico e finalistico laddove il primo impiega pochi e selezionati algoritmi continuamente aggiornati mentre il secondo consente un'ampia scelta di utilizzo. Una differenza sostanziale sul piano tecnico tra le due librerie consiste nell'oggetto di elaborazione; NLTK è orientata verso il processamento di stringhe mentre spaCy verso gli oggetti. Trattandosi di estensioni con specifiche proprietà e metodi, le stringhe sono propense ad un maggiore grado di aggiornamento comportando, nel caso di NLTK, di dover controllare periodicamente la documentazione per scoprire eventuali nuove funzioni.

GENSIM. Un'altra libreria open source di largo impiego è GENSIM, che fa uso di algoritmi non supervisionati quali *Word2Vec*, *FastText*, *LdaModel* etc. per rintracciare all'interno di corpora testuali pattern statistici da cui partire per elaborare la modellazione di argomenti (dall'inglese Topic modelling) dei documenti. Questa libreria è stata sviluppata per individuare i topic su grandi corpus di testi facendo ricorso a due principali metodi: *Latent Semantic Analysis* (LSA) e *Latent Dirichlet Allocation* (LDA). LSA è una tecnica di *Information Retrieval* (IR) impiegata per analizzare le relazioni tra un set di documenti e i termini in essi contenuti con l'obiettivo di estrarre i concetti assunti dalle parole nei documenti. LDA, come descritto in 3.1, è un modello probabilistico generativo dei topic basato sulla frequenza delle parole in un set di documenti che necessita di un corpus già tokenizzato e lemmatizzato per funzionare. Per questa fase di pre-processing GENSIM impiega NLTK per generare, successivamente, bigrammi (set di due parole adiacenti) e Bag-of-words sulla base della frequenza delle parole nei documenti.

BERT. Bidirectional Encoder Representations from Transformers, è un modello di ML basato sui Transformer, un meccanismo di attenzione che apprende le relazioni contestuali tra le parole in un testo. BERT è stato sviluppato con l'obiettivo di colmare la necessità di grandi dati di addestramento annotati di cui l'NLP necessita per essere accurato. Per ottemperare a questa mancanza, BERT è composto di due fasi principali: pre-training e fine-tuning. Nella fase di pre-addestramento il framework è addestrato su dati non annotati impiegando due processi non supervisionati: *Masked LM* (MLM) e *Next Sentence Prediction* (NSP). MLM è un modello bidirezionale per la predizione di sequenze di token addestrato attraverso il riconoscimento di token "mascherati"³⁵¹, presenti in una certa percentuale all'interno di singole frasi. Questo modello ha la funzione di predire esclusivamente le parole mascherate anziché ricostruire l'intero segmento testuale in input. NSP è un processo

³⁵¹ In letteratura il termine impiegato è *masked*, da qui il nome del modello Masked LM.

finalizzato a rendere il modello capace di comprendere la relazione tra due frasi e può essere generato da qualsiasi corpus monolingua.

Nella fase di *fine-tuning*, i parametri vengono ottimizzati con l'aggiunta di specifici dati annotati in base alla tipologia di processo da effettuare, come illustrato nella Figura 40. Questa fase consente di perfezionare il modello sulla base alla specifica attività, ottenendo risultati rilevanti senza dover ricorrere a grandi quantità di dati annotati.

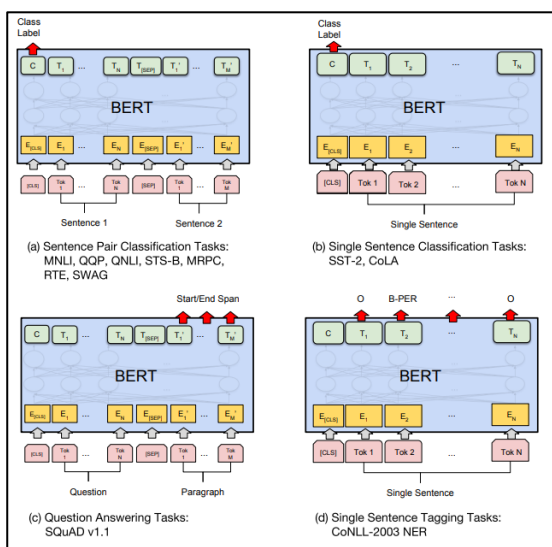


Figura 40: Fase di Fine-tuning di BERT su diverse elaborazioni³⁵².

Sulla base di BERT, sono nati diversi modelli per l'elaborazione di specifici task come, ad esempio:

- BERTopic, sviluppato per generare le rappresentazioni dei topic attraverso tecniche di clustering e variazione su classi di Term Frequency – Inverse Document Frequency (TF-IDF). Per raggiungere l'obiettivo finale, il modello è implementato su tre fasi³⁵³. Nella prima vengono creati delle rappresentazioni vettoriali dei documenti (*document embeddings*) applicando modelli linguistici pre-addestrati per il recupero delle informazioni a livello di documento. Successivamente, dopo aver ridotto le rappresentazioni vettoriali, si procede alla creazione di cluster semantica di documenti rappresentanti i diversi topic. Infine, nella fase conclusiva, viene applicato il TF-IDF basato su classi per estrarre la rappresentazione finale degli argomenti.
- AIBERTO³⁵⁴, basato su BERT per l'elaborazione di processi di NLP addestrato specificatamente sulla lingua italiana e, in particolare, su testi provenienti dai social network.

³⁵² Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding" *cit.* p.15.

³⁵³ Maarten Grootendorst, "BERTopic: Neural topic modeling with a class-based TF-IDF procedure," (2022), p. 2. <https://arxiv.org/pdf/2203.05794.pdf>

³⁵⁴ Lan Zhenzhong et al., "ALBERT: A Lite BERT for Self-supervised Learning of Language Representations," (2020). <https://arxiv.org/abs/1909.11942>.

- RoBERTa³⁵⁵, un modello che si pone come un “avanzamento” di BERT sotto il profilo dell’addestramento. Secondo quanto affermato dagli autori³⁵⁶, il modello nasce per ovviare ad alcuni limiti mostrati da BERT in fase di addestramento. Nel dettaglio, le modifiche proposte dagli autori hanno riguardato una più lunga fase di addestramento e su frasi di lunghezza maggiore, la rimozione del processo di NSP, un cambio dinamico del mascheramento dei token nei dati di addestramento. Il modello, testato coi medesimi parametri di BERT, ha mostrato un netto miglioramento dimostrando come la concezione del pre-addestramento portata avanti da BERT rimanga competitiva.

Estrazione automatica di informazioni da grafici

La conoscenza contenuta nei grafici presenti in testi pubblicati, anche in report e statistiche turistiche non è di facile estrazione in quanto tali oggetti risultano inseriti nei testi sotto forma di immagine statica. Negli ultimi anni sono stati sviluppati diversi framework in grado di processare questi documenti col fine di rendere le informazioni contenuti *machine-readable*. Nell’attuale stato dell’arte si distinguono applicativi *web-based* come gli open source WebPlotDigitizer³⁵⁷ e GraphReader³⁵⁸ o il servizio a pagamento PlotDigitizer³⁵⁹, e applicativi stand-alone come ReVision³⁶⁰, sviluppato in codice MathLab, Scatteract³⁶¹, ChartReader³⁶² e ChartOCR³⁶³ sviluppati in Python. La prima categoria si basa su un approccio prevalentemente manuale, in cui è compito dell’utente delineare area e oggetti di interesse all’interno del grafo. I framework afferenti alla seconda categoria invece, implementano al loro interno complessi algoritmi di machine e deep learning in grado di analizzare ed estrarre la diversa componente informativa presente in queste risorse. Per la definizione di tali applicativi l’approccio generale, come ad esempio in (Liu, Klabjan, and Bless 2019)³⁶⁴, prevede la definizione di un workflow così suddiviso:

1. Riconoscimento della tipologia di grafico rappresentato;

³⁵⁵ Liu Yinhan et al., “RoBERTa: A Robustly Optimized BERT Pretraining Approach,” (2019). arxiv.org/pdf/1907.11692.

³⁵⁶ Liu Yinhan et al., “RoBERTa: A Robustly Optimized BERT Pretraining Approach,” *cit.* p.1.

³⁵⁷ Ankit Rohatgi, “WebPlotDigitizer,” ultimo accesso 23 Gennaio 2024, <https://automeris.io/WebPlotDigitizer/>.

³⁵⁸ Kristian Pontoppidan, “GraphReader,” ultimo accesso 23 Gennaio 2024, <http://www.graphreader.com/>,

³⁵⁹ Porbital, “PlotDigitize,” ultimo accesso 23 Gennaio 2024, <https://plotdigitizer.com/>.

³⁶⁰ Manolis Savva et al., “ReVision: Automated Classification, Analysis and Redesign of Chart Images,” *ACM User Interface Software & Technology (UIST)*, (2011). DOI:10.1145/2047196.2047247.

³⁶¹ Mathieu Cliche, David Rosenberg, Dhruv Madeka, and Connie Yee. “Scatteract: Automated extraction of data from scatter plots,” (2017). arXiv:1704.06687.

³⁶² Chinmayee Rane et al., “ChartReader: Automatic Parsing of Bar-Plots,” *2021 IEEE 22nd International Conference on Information Reuse and Integration for Data Science (IRI)*, (2021). DOI:10.1109/IRI51335.2021.00050.

³⁶³ Junyu Luo, Zekun Li, Jinpeng Wang, and Chin-Yew Lin, “ChartOCR: Data Extraction from Charts Images via a Deep Hybrid Framework,” *2021 IEEE Winter Conference on Applications of Computer Vision (WACV)*. 2021. https://openaccess.thecvf.com/content/WACV2021/papers/Luo_ChartOCR_Data_Extraction_From_Charts_Images_via_a_Deep_Hybrid_WACV_2021_paper.pdf

³⁶⁴ Xiaoyi Liu, Diego Klabjan, and Patrick N. Bless. “Data extraction from Charts via Single Deep Neural Network,” (2019). arXiv:1906.11906.

2. individuazione degli assi x e y che compongono il grafico e delimitazione dell'area di interesse;
3. riconoscimento ed estrazione della componente testuale (etichette, legenda, valori, etc.);
4. calcolo dei valori numerici basato sulla distanza fra pixel di oggetti individuati nell'area;
5. definizione di regole per la conversione dei dati in formato *machine-readable*.

Relativamente al primo punto, occorre tenere in considerazione la varietà di grafici esistenti per i quali è richiesta la costruzione di specifiche regole per la corretta individuazione degli elementi.

Per la sperimentazione di tale processo di estrazione automatica è stata presa in esame la categoria di strumenti di estrazione automatica. A seguito di un'attività di verifica e controllo della disponibilità degli applicativi sopra citati, per la sperimentazione nel presente progetto di dottorato sono stati presi in esame gli unici realmente disponibili in Python: ChartOCR³⁶⁵ e ChartReader³⁶⁶. La tabella sottostante offre una sintesi dei due framework.

Name	ChartOCR	ChartReader
Programming Language	Python 3.6	Python 3.8
Code Availability	Yes	Yes
Grafici	Completo	Barre ³⁶⁷
Issue(s)	Compute power too expensive Not easy to relate all packages	Under testing
Operating System	Linux Ubuntu	Windows
Facilities	Use Anaconda Environment	Jupyter Notebooks/Google Colab

Tabella 9: Comparazione tra Chart OCR e ChartReader.

Sulla base di una siffatta constatazione si è stabilito di applicare ChartReader³⁶⁸ per l'estrazione delle informazioni dai grafici a barre. A differenza di ChartOCR, tale framework presenta un'architettura

³⁶⁵ Junyu Luo, "DeepRule," Ultimo accesso 23 Gennaio 2024, <https://github.com/soap117/DeepRule>.

³⁶⁶ Chinmayee Rane, "ChartReader," ultimo accesso 23 Gennaio 2024, <https://github.com/Cvrane/ChartReader..>

³⁶⁷ Il codice sorgente attualmente disponibile include la sola tipologia di grafico a barra per una estrazione totale e accurata.

³⁶⁸ Il framework è stato ulteriormente testato all'interno di un lavoro di ricerca internazionale a più mani per l'estrazione automatica della conoscenza in un dominio diverso e pubblicato in Anna Rovella, Alexander Murzaku, Eugenio Cesario, Martin Critelli, Armando Bartucci, and Francesca M. C. Messiniti, "Analysis, evaluation and comparison of knowledge extraction tools in the Environmental and Health domain. A holistic approach," *Knowledge Organization and Management in the Domain of Environment and Earth Observation (KOMEEO)*, (Ergon-Verlag, 2022). DOI:10.5771/9783956508752-121.

meno complessa e, come tale, non richiede una eccessiva potenza di calcolo. ChartReader è composto di diversi moduli che consentono di:

1. estrarre elementi di contesto del grafico;
2. riconoscere la tipologia di grafico e gli assi attraverso due modelli di CNN: EfficientNet e VGG-19;
3. estrarre il testo attraverso le API di Amazon Web Services;
4. rappresentare i dati estratti.

In questo caso di studio, il framework è stato testato su un piccolo campione di grafici statistici relativi al dominio del turismo culturale.

*Alla Ricerca, colei che mi ha affascinato,
entusiasmato, disilluso e insegnato tanto.*