



The banner features a row of six icons: a globe, a book, a handshake, a money bag with a Euro symbol, a scale of justice, and a bicycle. Below the icons, the text reads: **AIUCD 2021**, **DH per la società: e-guaglianza, partecipazione, diritti e valori nell'era digitale**, and **10° congresso annuale PISA 19-22 gennaio**. On the right side, a list of topics is displayed in colored text: **DIGITAL PUBLIC HUMANITIES** (red), **OPEN CULTURE** (orange), **RETI SOCIALI** (yellow), **TECH ECONOMY** (green), **E-PARTICIPATION** (blue), and **TECNOLOGIE ASSISTIVE** (purple). The background includes binary code and a classical building facade.

Versione PROVVISORIA del contributo presentato al Convegno Annuale

DISCLAIMER

Questa versione dell'abstract non è da considerarsi definitiva e viene pubblicata esclusivamente per facilitare la partecipazione del pubblico al convegno AIUCD 2021

Il Book of Abstract contenente le versioni definitive e dotato di ISBN sarà disponibile liberamente a partire dal 19 gennaio sul sito del convegno sotto licenza creative commons.

Verso la descrizione automatica delle immagini nell'editoria digitale accessibile: proposta di una tassonomia di immagini per gli algoritmi di intelligenza artificiale

Chiara De Martin¹, Barbara Leporini², Gregorio Pellegrino³

¹Università di Pisa, Italia- c.demartin1(«»)studenti.unipi.it

²ISTI-CNR, Italia- barbara.leporini(«»)isti.cnr.it

³Fondazione LIA, Italia- gregorio.pellegrino(«»)fondazionelia.org

ABSTRACT

In questo contributo viene proposta una tassonomia di possibili tipologie di immagini da poter utilizzare per migliorare la composizione dei dataset di addestramento degli algoritmi di intelligenza artificiale da applicare ai task di classificazione e descrizione automatica delle immagini. Viene altresì sinteticamente illustrato il processo di validazione effettuato nello studio. 25 categorie sono state identificate nella tassonomia proposta.

PAROLE CHIAVE

Editoria digitale accessibile; accessibilità immagini; descrizione alternativa immagini; tassonomia immagini; intelligenza artificiale; Fondazione LIA

1. INTRODUZIONE

Il problema dell'accesso ai contenuti non testuali è molto sentito nella comunità delle persone con disabilità. L'accesso alle immagini da parte di utenti che usano uno screen reader è stato risolto dalle linee guida sull'accessibilità con l'adozione di descrizioni alternative del contenuto degli oggetti non testuali, che rappresentano un requisito fondamentale per l'accessibilità di qualunque pubblicazione digitale. Descrivere le immagini è un'attività qualificata, costosa e lunga e si configura come uno dei requisiti più difficile da soddisfare. Una possibile soluzione è quella di automatizzare o semi-automatizzare il processo tramite l'intelligenza artificiale [7]. Pertanto, proponiamo una tassonomia per la classificazione delle immagini, con lo scopo di aiutare gli algoritmi a produrre descrizioni efficaci in base al tipo di immagine.

Per ogni categoria della tassonomia vengono proposti degli attributi che ne descrivano le possibili caratteristiche di forma e di contenuto.

2. METODOLOGIA

Per definire la tassonomia sono state consultate le principali linee guida internazionali sulla descrizione alternativa delle immagini [2], [5] e [6]. Le categorie da queste analizzate sono state combinate con quelle emerse dall'esame di metodi manuali, semi-automatici e automatici per la descrizione alternativa delle immagini. Sono stati esaminati allo scopo anche tool e strumenti online dedicati al task, come [3], banche dati di immagini¹ e dataset gratuiti². La tassonomia è stata ottenuta combinando le categorie prese in considerazione dalle diverse fonti e basandosi sulle caratteristiche associate a ciascuna di esse per definire i relativi attributi e i rispettivi possibili valori.

La tassonomia è stata validata su alcuni manuali di editoria scolastica; la metodologia applicata è stata di tipo qualitativo. La validazione è stata condotta cercando di collocare ogni immagine contenuta nei manuali in una categoria e di descriverla mediante gli attributi corrispondenti. Nuove categorie e attributi sono stati aggiunti per colmare lacune e inadeguatezze.

¹ iStock by Getty Images: <<https://www.istockphoto.com/it/>>; Pixabay: <<https://pixabay.com/it/>>; Internet Archive, "Community Images": <https://archive.org/details/opensource_image/>; Flickr: <<https://flickr.com/>>.

² MSCOCO: <<http://cocodataset.org/>>; Flickr8k Dataset: <<http://academic.toronto.edu/~tfliu/flickr8k/>>; Flickr30k: <<https://www.kaggle.com/hsankesara/flickr-image-dataset>>; PASCAL: <<http://host.robots.ox.ac.uk/pascal/VOC/>>.

3. PROPOSTA DI UNA TASSONOMIA

La tassonomia prevede venticinque categorie: fotografie, disegni, dipinti, fumetti/comic, cartoon, mappe, grafici, diagrammi, tabelle, timeline, cover, icone, loghi, bandiere, firme, testo / testo riempito con immagini, typeface, screenshot, poster, immagini complesse / infografica / organigrammi, meme, segnaletica, pubblicità, collage, murali. Alcune categorie, come grafici o diagrammi, comprendono sottocategorie (grafici a barre, a linea, a torta, ...).

Per ogni categoria sono proposti degli attributi, formulati come campi per i quali vengono previsti una serie di possibili valori. La selezione degli attributi è stata guidata dall'obiettivo di catturare nel modo più dettagliato possibile gli aspetti rilevanti che possono contraddistinguere le immagini delle varie categorie. Alcuni attributi sono sprovvisti di valori predefiniti, poiché la loro variabilità è tale da rendere impossibile fornire a priori delle indicazioni. Alcuni attributi sono di carattere generale e applicabili a più di una tipologia di immagini; altri sono più specifici e hanno un'applicabilità più ristretta. Alcuni valori svolgono a loro volta il ruolo di attributi: in questo caso, il loro valore è rappresentato da un'altra etichetta, o da una combinazione di più etichette.

Nella tassonomia sono inclusi due attributi di validità globale: "rilevanza dell'immagine" e "tipo di descrizione". Il primo serve per tenere conto del fatto che un'immagine può essere, nel contesto in cui è inserita, *significativa* o *non significativa* e dunque avere bisogno o meno di essere descritta; il secondo specifica il tipo di descrizione più adatto all'immagine: questa può essere breve o lunga, ma composta di solo testo, oppure lunga e strutturata, richiedendo una formattazione che prevede l'uso di liste o tabelle.

4. VALIDAZIONE

La tassonomia è stata validata su quattro libri di testo scolastici, di materie diverse: biologia, comunicazione, tecnologia, storia. Per ogni immagine, si è cercato di classificarla utilizzando le categorie individuate e di descriverla usando i campi degli attributi. Nuove categorie sono state aggiunte per coprire immagini che altrimenti non avrebbero trovato un'etichetta adeguata (come "emoticon" o "francobollo"); nuovi attributi sono stati aggiunti per raffinare le categorie esistenti.

Il riscontro ottenuto in fase di validazione si è rivelato un ottimo feedback per valutare possibili modifiche future.

5. CONCLUSIONI

In questo contributo si propone una possibile tassonomia per il raffinamento della composizione dei dataset per l'addestramento degli algoritmi di intelligenza artificiale, con l'obiettivo di contribuire a un'editoria digitale sempre più accessibile, con particolare riferimento al tema della descrizione alternativa delle immagini. La tassonomia proposta include venticinque categorie. La validazione condotta ha dimostrato come la tassonomia sia in grado di descrivere un'elevata percentuale di immagini ed elementi non testuali. Si ritiene che la stessa possa costituire un valido punto di riferimento per la composizione dei dataset, portando in questo modo a un miglioramento delle prestazioni e dell'accuratezza degli algoritmi nei task di classificazione e descrizione automatica delle immagini.

6. RINGRAZIAMENTI

Si ringrazia l'editore Zanichelli per aver messo a disposizione il suo catalogo digitale per validare la tassonomia.

BIBLIOGRAFIA

- [1] Amornchat, Supada. «Complex Images for All Learners», s.d. <https://www.pcc.edu/instructional-support/wp-content/uploads/sites/17/2017/11/complex-images.pdf>.
- [2] DIAGRAM Center. «Image Description Guidelines». DIAGRAM Center, giugno 2015. <http://diagramcenter.org/table-of-contents-2.html>.
- [3] DIAGRAM Center. «Poet training tool». <https://poet.diagramcenter.org/>
- [4] Gleason, Cole, Amy Pavel, Xingyu Liu, Patrick Carrington, Lydia B. Chilton, e Jeffrey P. Bigham. «Making Memes Accessible». In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility*, 367–376. ASSETS '19. New York, NY, USA: Association for Computing Machinery, 2019. <https://doi.org/10.1145/3308561.3353792>.
- [5] NCAM. «Effective Practices for Description of Science Content within Digital Talking Books». WGBH, dicembre 2008. <https://www.wgbh.org/foundation/ncam/guidelines/effective-practices-for-description-of-science-content-within-digital-talking-books>.
- [6] NWEA. «NWEA Image Description Guidelines for Assessments». Northwest Evaluation Association, 2016. https://www.nwea.org/content/uploads/2016/10/Image-Description-Guidelines-IDGA_June2016.pdf.
- [7] Pellegrino, Gregorio. «E-book. Automatizzare le descrizioni alternative delle immagini con l'AI». *Giornale della Libreria*, 22 luglio 2019. <https://www.giornaledellalibreria.it/news-innovazione-e-book-automatizzare-le-descrizioni-alternative-delle-immagini-con-lai-3877.html>.

Data di ultimo accesso degli URL: 30 novembre 2020