

Linee-guida annotazione

Introduzione

I testi da annotare fanno parte di un corpus di testi utile ad istruire un classificatore automatico per la classificazione di contenuti multimediali online, per comprendere se questi siano adatti o meno ai minori e quale sia l'età giusta per fruire di un determinato contenuto.

Il corpus è formato da trascrizioni di video, canzoni e testo generato online dagli utenti (in inglese e in italiano). Il compito di annotazione consiste nell'annotare l'età adatta per fruire di quel determinato contenuto, in base alla propria esperienza di parlante (per tema e complessità del linguaggio) e in base alla direttiva AGCOM sulla classificazione delle opere audiovisive destinate al web, e nell'annotare la presenza o meno di parole che fanno riferimento ai descrittori tematici (discriminazione, droghe, nudità, sesso, minacce, violenza).

Interfaccia di Annotazione

Per le annotazioni, ti servirai di un foglio di calcolo su Google Sheets. I testi da annotare sono raccolti in una cartella che verrà condivisa con gli annotatori su Google Drive, ogni file di testo ha un titolo che corrisponde al titolo riportato nella prima colonna del file di calcolo, fai molta attenzione alla corretta corrispondenza tra titolo del file e riga sulla quale stai lavorando. Puoi sfruttare il collegamento ipertestuale di ogni titolo per accedere direttamente al testo di riferimento nella cartella. Alla destra del titolo, puoi selezionare la categoria appropriata per ogni testo.

Nell'interfaccia di annotazione per ogni colonna c'è un menù a discesa delle possibili categorie (etichette): premendo "Invio" o facendo clic con il mouse sulla cella selezionata, visualizzerai un menu a discesa con tutte le possibili etichette, seleziona quella appropriata.

1	Titolo	Età	Discriminazione	Droghe	Comportamento peric	Linguaggio scurril	Nudità	Sesso	Minacce	Violenza	Complessità del linguaggio	ID Annotatore	Lingua
2	Ted2020-1											46,1_002_A	Italiano
3	Ted2020-2	3										46,44_002_A	Italiano
4	Ted2020-3											46,27_002_A	Italiano
5	testo_ytkids_0001.txt	6										002_A	Italiano
6	Ted2020-5	12										46,98_002_A	Italiano
7	testo_ytkids_0002.txt	15										002_A	Italiano
8	testo_ytkids_0003.txt											002_A	Italiano
9	Ted2020-8	18										45,64_002_A	Italiano
10	testo_ytkids_0011.txt											002_A	Italiano
11	Ted2020-10											35,09_002_A	Italiano

Figura 1. Esempio dell'interfaccia di annotazione.

Età di riferimento

I riferimenti per la categorizzazione sono ispirati al sistema PEGI per la classificazione dei videogiochi e alle linee guida dell'AGCOM sulla classificazione delle opere audiovisive destinate al web e sono i seguenti:

- **Opere per tutti (3):**
 - Il contenuto del testo è considerato adatto a tutte le età.
 - Il testo non deve contenere nessun linguaggio né comportamenti **discriminatori né incitamento all'odio**¹ se non in contesto di iniziative sociali.
 - Non deve esserci nessun riferimento a droghe o dipendenza da sostanze (droga, alcol, etc.)
 - Non può esserci nessun accenno a comportamenti pericolosi o facilmente imitabili.

¹ Per "discriminazione e incitamento all'odio" deve intendersi l'utilizzo di contenuti o espressioni che diffondono, propagandano l'odio e la discriminazione e istigano alla violenza nei confronti di un determinato target attraverso stereotipi relativi a caratteristiche di gruppo, etniche, di provenienza territoriale, di credo religioso, d'identità di genere, di orientamento sessuale, di disabilità, di condizioni personali e sociali conformemente alla delibera AGCOM n.15719/CONS.

- Il linguaggio non può essere scurrile.
- Non è consentita nessuna nudità integrale. Rappresentazioni sporadiche, occasionali e non insistenti di nudità parziale sono ammesse se riconducibili a contesti naturali, scientifici e/o di quotidianità, e in ogni caso non riconducibili in alcun modo a contesti sessuali.
- Deve esserci totale assenza di rappresentazione di attività sessuali. Rappresentazioni affettive sono ammesse se limitate a effusioni non insistenti e totalmente prive di riferimenti erotici.
- I comportamenti violenti devono essere assenti ma sono ammessi se riconducibili a contesti non realistici (es. gag comiche)
- Il testo non deve contenere elementi (parole, frasi, riferimenti a scene, trascrizione di suoni) che potrebbero spaventare i bambini piccoli. Una forma di violenza molto lieve (in un contesto comico o in un ambiente infantile) è accettabile.
- **Opere non adatte ai minori di 6 anni (6):**
 - Il testo non deve contenere nessun linguaggio né comportamento discriminatorio a meno che (a) non sia chiaramente disapprovato o inserito in contesti storici, (b) abbia fini educativi.
 - La rappresentazione di droghe e dipendenza da sostanze (droga, alcol etc.) è ammessa esclusivamente con evidente finalità divulgativa ed educativa, o se collegata ad iniziative sociali (es. campagne di prevenzione, sensibilizzazione e informazione).
 - Può esserci una rappresentazione occasionale, accennata o irrealistica di comportamenti pericolosi, con evidente disapprovazione.
 - La presenza di turpiloquio e di termini volgari di uso comune nella quotidianità è consentita solo se sporadica.
 - Rappresentazioni di nudità parziale sono ammesse solo se non sono riconducibili a un contesto evidentemente sessuale.
 - Non possono esserci rappresentazioni di attività sessuali. Le effusioni affettive sono ammesse se non sono insistenti.
 - E' prevista la totale assenza di sequenze che possano indurre stati di agitazione e di tensione emotiva.
 - Può esserci presenza sporadica di sequenze idonee a indurre stati di agitazione e di tensione emotiva. L'esito deve essere positivo.
 - Possono esserci scene di violenza blanda e di breve durata, riconducibili a generi finzionali codificati come di generalizzata fruizione da parte di minori, di cui non sono visibili le conseguenze (es. sangue e ferite) ovvero categorizzati dal fine educativo.
 - Possono essere presenti elementi potenzialmente impressionanti per i bambini.
 - Il testo potrebbe contenere qualche elemento che potrebbe spaventare i bambini molto piccoli. Rientra in questa categoria un testo contenente una forma di violenza molto lieve (implicita o non realistica).
- **Opere non adatte ai minori di 12 anni (12):**
 - Il testo può contenere un linguaggio e/o comportamenti discriminatori giustificati dal contesto narrativo a condizione che: (a) siano espressamente condannati e/o sanzionati moralmente, ovvero inseriti in contesti con chiari fini educativi o di rievocazione storica, (b) che siano riconducibili nell'immaginario collettivo a valori e/o personaggi negativi rispetto ai quali sia esclusa una possibile identificazione da parte dei minori.
 - Il testo non può contenere nessun incitamento all'odio.
 - La rappresentazione di droghe e dipendenza da sostanze (droga, alcol etc.) è consentita a condizione che (a) sia collegata a contesti divulgativi e/o educativi, (b) sia accompagnata da messaggi di chiara condanna.
 - La presenza di comportamenti pericolosi facilmente imitabili può esserci, a condizione che non siano presentati come affascinanti, non siano sollecitati e vi sia un messaggio finale positivo.

- La presenza, non reiterata, gratuita o insistita di termini volgari e il ricorso al turpiloquio sono consentiti solo se inseriti in contesti idonei a stemperare potenziali criticità (es. contesto comico).
 - Nudità integrali occasionali e non insistite sono ammesse se sono avulse da contesti sessuali. Rappresentazioni sporadiche, o accennate, di nudità parziali sono ammesse anche in contesti sessuali.
 - Possono esserci rappresentazioni di attività sessuali poco frequenti e limitate effusioni e prive di linguaggi espliciti e provocatoriamente allusivi.
 - Possono essere presenti sequenze spaventose, e/o di pericolo, idonee a indurre stati di agitazione e di tensione solo se (a) di breve durata, (b) con esito finale giustificabile in base agli strumenti di comprensione della relativa fascia di età.
 - E' ammessa la presenza di scene di blanda violenza solo se (a) contestualizzata, (b) sanzionata moralmente, (c) priva di insistenza morbosa su particolari crudi o brutali, rappresentabili solo se giustificati dal contesto.
 - Possono essere presenti sequenze horror o rappresentazioni raccapriccianti.
 - Possono esserci elementi che si riferiscono all'utilizzo di giochi di possibilità, rappresentati come tipici giochi d'azzardo che risultino di importanza secondaria rispetto all'intera esperienza di gioco.
 - Rientrano in questa categoria i testi che fanno riferimento a violenza di natura leggermente più esplicita verso personaggi fantasy o violenza non realistica nei confronti di personaggi umanizzati.
- **Opere non adatte ai minori di 15 anni (15):**
 - Il testo può contenere linguaggio e/o comportamenti discriminatori, purché non rappresentati in chiave positiva e non incoraggiati.
 - L'incitamento all'odio è ammesso solo in presenza di chiara condanna.
 - La rappresentazione di droghe e dipendenza da sostanze (droga, alcol, etc.) è ammessa purché non in chiave positiva e priva di legittimazione e/o esaltazione.
 - La presenza di comportamenti pericolosi facilmente imitabili può essere presente frequentemente ma devono essere scoraggiati o rappresentati come difficilmente imitabili.
 - La presenza frequente di termini volgari e/o offensivi ed il ricorso al turpiloquio sono consentiti solo se (a) non vengono rappresentati come standard positivi, (b) vengono sanzionati moralmente.
 - Nudità integrali sono ammesse se avulse da contesti sessuali e non insistite. Rappresentazioni, anche frequenti, di nudità parziali sono ammesse anche in contesti erotici.
 - Rappresentazioni di attività sessuali/erotiche sono ammesse anche se accompagnate da linguaggi espliciti e provocatoriamente allusivi, purché poco frequenti e prive di dettagli insistiti e ravvicinati. Riferimenti a violenza sessuale possono essere presenti, in un rapporto di funzionalità rispetto alla narrazione, ma senza insistenza e di breve durata.
 - Possono essere presenti sequenze idonee a indurre stati di agitazione e di alta tensione emotiva purché di durata non eccessiva rispetto al contesto e prive di insistenza morbosa su particolari crudi o brutali.
 - La presenza di scene di violenza, anche insistita ed efferata, è ammessa solo se (a) non rappresentata in chiave positiva, (b) priva di insistenza morbosa su particolari crudi o brutali, rappresentabili solo se giustificati dal contesto.
 - Possono essere presenti sequenze horror o rappresentazioni raccapriccianti.
 - Possono esserci elementi che si riferiscono all'utilizzo di giochi di possibilità giocati/rappresentati come tipici giochi di azzardo, che risultino di importanza secondaria rispetto all'intera esperienza di gioco.

- **Opere non adatte ai minori di 18 anni (18):**

- Il testo può contenere libera rappresentazione di comportamenti intolleranti, razzisti o sessisti.
- L'incitamento all'odio, da parte di personaggi negativi, deve essere ricomposto e riequilibrato nella narrazione.
- Libera rappresentazione dell'uso, del reperimento e della fabbricazione di droghe.
- Libera rappresentazione di comportamenti pericolosi.
- La presenza molto frequente di termini volgari e/o offensivi ed il ricorso al turpiloquio sono consentiti. Esclusa la blasfemia non occasionale o ripetuta.
- Le nudità integrali sono ammesse anche in contesti sessuali e/o erotici purché non oggetto di scene pornografiche.
- Possono essere presenti rappresentazioni di attività sessuali ed erotiche anche accompagnate da linguaggi espliciti, giustificate dalla trama e dal contesto.
- Libera presenza di sequenze angosciose o spaventose ad elevato impatto emotivo.
- La presenza di scene esplicite di violenza, anche gratuita, insistita ed efferata, è ammessa solo se non legittimata e/o esaltata. Esclusa la rappresentazione di violenza sadica.
- Possono essere presenti sequenze horror o rappresentazioni raccapriccianti.
- Possono esserci elementi che si riferiscono all'utilizzo di giochi di possibilità, giocati/rappresentati come tipici giochi d'azzardo, che risultino di importanza secondaria rispetto all'intera sequenza di gioco.

Identificazione della presenza di parole che fanno riferimento ai descrittori tematici

Le colonne "Discriminazione", "Droghe", "Comportamento pericoloso", "Linguaggio scurrile", "Nudità", "Sesso", "Minacce", "Violenza" consentono, attraverso un menù a tendina che comprende due scelte (Sì/No), di annotare la presenza o meno di parole o frasi all'interno del testo, che fanno riferimento a quella specifica categoria.

Ad esempio: il testo dal titolo "*Ted2020-171*" contiene le seguenti frasi:

"E tu che cazzo ne sai della Guardia Nazionale?", "Figli di puttana!", "Se uno di quei medici incompetenti mi avesse detto di interrompere la cura, gli avrei tagliato la gola sul posto."

Quindi nella riga relativa sceglierai **Sì** nella colonna del *Linguaggio scurrile*, in quella delle *Minacce* e in quella della *Violenza*, e se il resto del testo non contiene riferimento agli altri indicatori annoteremo **No** nel resto delle caselle.

1	Titolo	Età	Discriminazione	Droghe	Comportamento pericoloso	Linguaggio scurrile	Nudità	Sesso	Minacce	Violenza
148	Ted2020-171	18 ▼	No ▼	No ▼	No ▼	Sì ▼	No ▼	No ▼	Sì ▼	Sì ▼

Figura 2. Esempio annotazione della presenza di parole che fanno riferimento agli indicatori tematici

Altre colonne: lingua, id annotatore, note

La colonna "Complessità del linguaggio" è riservata ad un'annotazione automatica e va lasciata vuota. Le colonne "Lingua" ed "ID Annotatore" sono già compilate. "Lingua" si riferisce alla lingua del testo da annotare che può essere "Inglese" o "Italiano". "ID Annotatore" è un riferimento che identifica chi ha annotato il testo. La colonna "Note" è una colonna riservata ai commenti.

Descrittori tematici delle opere audiovisive destinate al web

Discriminazione e incitamento all'odio	Utilizzo di contenuti o espressioni che diffondono, propagandano l'odio e la discriminazione e istigano alla violenza nei confronti di un determinato target attraverso stereotipi relativi a caratteristiche di gruppo, etniche, di provenienza territoriale, di credo religioso, d'identità di genere, di orientamento sessuale, di disabilità, di condizioni personali e sociali conformemente alla delibera AGCOM n.15719/CONS.
Droghe	
Comportamento pericoloso e facilmente imitabile	
Linguaggio	
Nudità	
Sesso	
Minacce	Per "minacce" deve intendersi la presenza di contenuti suscettibili di indurre nello spettatore stati di tensione emotiva, paura e angoscia derivati dalle dinamiche strutturali del racconto (es. contrapposizione bene/male); dilemmi etici; dinamiche conflittuali tra i personaggi; quadro valoriale sotteso alla storia).
Violenza	

Suggerimenti utili

1. Fare **molta attenzione**, quando si annotano i TED Talks, o comunque testi lunghi che non possono essere contenuti nel file di Google Sheet stesso, che il titolo del testo che si sta leggendo corrisponda esattamente al titolo riportato nella colonna "Titolo" del file di Google Sheet.
2. È buona norma scrivere un report durante l'annotazione sul quale annotare eventuali dubbi, perplessità o errori riscontrati durante il lavoro, e inviarlo periodicamente per e-mail a apaone@unior.it.
3. Per non compromettere la validità del lavoro non consultarti con gli altri annotatori riguardo gli stessi testi.
4. Non aprire eventuali link ad altri siti presenti nel testo.
5. Non modificare gli altri campi del file di Google Sheet senza chiedere, anche se si dovessero rilevare eventuali errori. Va comunque tutto segnalato.
6. Se il testo è scritto in una lingua diversa dall'italiano o dall'inglese, segnalalo.
7. Il testo potrebbe contenere errori o ripetizioni, si tratta di errori dovuti alla trascrizione automatica, vanno ignorati in quanto non compromettono il lavoro in questa fase.
8. In caso di dubbi o perplessità, non esitare a contattarmi.