

Il “giudice” algoritmo di fronte al fenomeno religioso

di Simone Baldetti

Abstract: The algorithm “judge” faces the religious phenomenon – The paper aims to highlight possible issues related to judicial applications of Artificial Intelligence and religion, starting from the specific case of content moderation in social media.

Keywords: Religion; AI; Social media; Internet; Hate speech.

1. Diritto e religione nello spazio digitale

Nel corso dei lavori di questo Convegno, abbiamo avuto occasione di riflettere su alcuni aspetti sociali legati all’uso delle moderne tecnologie in materia di Intelligenza Artificiale (IA). Con questo intervento intendo segnalare alcuni profili relativi al fenomeno religioso.

In linea generale, le religioni hanno mostrato di adattarsi piuttosto bene alle conseguenze sociali dello sviluppo delle tecnologie digitali, in particolar modo rispetto ad Internet, che per esse costituisce un formidabile strumento di comunicazione. Lo spazio digitale, esattamente come quello fisico, si è ben presto arricchito della presenza pubblica della religione, in forma tanto istituzionale – pensiamo agli account social network del Papa – quanto sociale, ossia legata al suo utilizzo come mezzo di divulgazione di discussioni religiose, e talvolta anche di propaganda. Le religioni si sono organizzate costituendo diverse tipologie di comunità virtuali: siti dedicati, forum, blog, gruppi di discussione nei social network, eccetera, che rivelano anche aspetti di originalità rispetto alle altre forme di interazione via web¹. Inoltre, un po’ tutti i soggetti religiosi collettivi si sono dotati di propri siti, a cominciare da quelli istituzionali (Conferenze episcopali, diocesi, parrocchie, eccetera) per finire con quelli – per così dire – di base: la presenza online della religione è ormai consolidata.

Il web ha anche ospitato riti religiosi, che si sono moltiplicati durante la pandemia da Covid-19, arrivando a sostituire le forme tradizionali in presenza², e

¹ Sia consentito il riferimento a S. Baldetti, *Il razzismo in Rete a 80 anni dalle leggi razziste di San Rossore. La tutela del sentimento religioso nelle regole delle online communities*, in P. Consorti (a cura di), *Costituzione, religione e cambiamenti nel diritto e nella società*, Pisa, 2019, 231 ss.; M. d’Arienzo, *Zuckerberg e i nuovi rapporti tra diritto e religioni. A proposito di libertà di coscienza nell’era digitale*, in *Diritto e Religioni*, n. 1-2019, 384-396.

² Sul tema delle conseguenze giuridico religiose della pandemia, vedi i contributi presenti nel volume di P. Consorti (a cura di), *Law, Religion and Covid-19 Emergency*, Pisa, 2020.

dando corpo a discussioni sulla possibile qualificazione dello spazio digitale come vero e proprio spazio liturgico³. Il propagarsi di questa presenza online impone di domandarsi anche in quale misura gli attuali sistemi di IA possono interfacciarsi col fenomeno religioso.

2. Hate speech online, religione e IA

Un esempio specifico della relazione tra IA e religione è quello relativo al ruolo dell'intelligenza artificiale come “giudice” rispetto ad aspetti della *governance* di Internet. Come è stato evidenziato dalla relazione della Prof.ssa Stradella, uno degli aspetti problematici dell'uso dell'IA in attività di tipo giudiziale, consiste nella possibilità di decisioni basate su criteri sì razionali, ma che finiscono per essere potenzialmente foriere di discriminazioni per particolari soggetti o gruppi.

Con riferimento alla regolamentazione dell'utilizzo dei moderni social media, è noto come la violazione delle regole di condotta delle piattaforme di discussione si manifesti anche attraverso comportamenti discriminatori nei confronti di minoranze religiose e culturali, sia attraverso la diffusione di fake news, sia in modo più diretto, mediante fenomeni di intolleranza e di vera e propria persecuzione. Con il passaggio al cosiddetto *web 2.0*, l'attività su internet ha visto un notevole incremento, dando vita ad un enorme numero di messaggi, immagini, video ecc. prodotti dagli utenti e trasmessi online attraverso mezzi come Facebook o Twitter⁴, tanto che è divenuto quasi impossibile poter eseguire dei controlli sul loro contenuto da parte di operatori umani. Per ciò si comincia ad utilizzare la capacità di calcolo dell'IA per controllare la compatibilità del contenuto messo sul web dall'utente con le regole delle piattaforme di social media. L'algoritmo diventa un vero e proprio “giudice” digitale che si sostituisce a quello umano nel vaglio di compatibilità tra i contenuti online e le regole di *governance* delle comunità virtuali⁵. Il risultato del giudizio ha però delle potenziali ricadute anche rispetto all'ordinamento giuridico civile, poiché alcuni dei comportamenti sanzionati hanno rilevanza per il diritto statale, per esempio in materia penale come nel caso della diffamazione. Per ciò un mancato intervento su un contenuto offensivo o discriminatorio può causare una vera e propria violazione di diritti fondamentali.

In questa prospettiva, affidare ad un algoritmo questo tipo di giudizio comporta far valutare all'IA, da un lato, il dato religioso come un indicatore di appartenenza, tutelando la scelta di appartenere ad un certo gruppo religioso o culturale e il conseguente diritto di non essere discriminati. Dall'altro, è necessario valutare in concreto l'esercizio del diritto di libertà religiosa nello spazio digitale, per esempio nell'utilizzo dell'elemento religioso ai fini di propaganda o discussione.

³ Cfr. L. M. Guzzo, *Il web può essere uno spazio liturgico?*, in *Il Regno.it*, 08.04.2020.

⁴ Cfr. S. Quintarelli, *Content moderation. I rimedi tecnici*, in G. Pitruzzella, O. Pollicino, S. Quintarelli, *Parole e potere. Libertà di espressione, Hate speech e Fake news*, Edizione Kindle, 2017.

⁵ Il cui funzionamento presenta alcune interessanti assonanze con quello delle *latae sententiae* canoniche, come nota R. Pascali, *Le leggi informatiche e le latae sententiae*, in *Diritto e Religioni*, n. 2-2011, 35 ss.

Il fenomeno religioso ha anche un ruolo significativo di contributo alla costruzione delle identità culturali, costituendo parte del background culturale dei componenti delle comunità. La complessità degli intrecci tra religione, diritto e cultura nelle moderne società multiculturali e la difficoltà di rispondere a tali istanze di tutela giuridica è un aspetto ormai noto⁶. In questa prospettiva, la sfida per il giudice algoritmo consiste nell'interpretare la complessità di un fenomeno legato ad aspetti fondamentali della persona umana. Per ciò, allorché si osservi la questione nei suoi aspetti più pragmatici, è possibile che le specificità del dato religioso rendano difficile per l'IA giudicare, aumentando così, ad esempio, la possibilità di verificarsi di falsi negativi rispetto alla valutazione dei contenuti online da parte dei sistemi algoritmici⁷.

Un esempio recente è il caso di un docente universitario dell'Università di Villanova, che si è visto sospendere temporaneamente il suo account su Facebook per un post contenente un'immagine presa dal film "Mission", che raffigurava una celebrazione religiosa cristiana. In questa immagine, un sacerdote guidava in cerimonia un gruppo di indios del Sud America, procedendo alla testa della processione con l'ostensorio⁸. Tuttavia, il contenuto era stato rimosso e la motivazione della sospensione consisteva nel fatto che una delle donne indios rappresentate nel film fosse a seno scoperto, ritenendo così violata la regola che prevedeva il divieto di rappresentazione di nudità nei post degli utenti⁹. In questo caso si rende necessario valutare la cornice della rappresentazione cinematografica utilizzata per veicolare un contenuto religioso.

Come è stato evidenziato da alcune ricerche, una delle maggiori sfide per lo sviluppo delle IA è quella di interpretare non tanto il singolo dato, piuttosto la complessità del contesto a cui è riferito un determinato contenuto¹⁰. La macchina non ha quindi la stessa sensibilità umana di capire che si trova davanti ad un post dal contenuto religioso in cui il focus è costituito da ciò che rappresenta la cerimonia cristiana rappresentata, mentre la circostanza della nudità costituisce un elemento del contesto (cinematografico) dell'immagine. Il semplice dato religioso non può quindi essere interpretato così com'è, ma va contestualizzato rispetto al messaggio che si vuole veicolare.

Pensiamo anche all'applicazione delle regole di derivazione religiosa. La diffusione di contenuti sul web che invocassero ad esempio il rispetto di determinati precetti coranici, per essere correttamente valutata richiederebbe una adeguata contestualizzazione della fattispecie. Questa esigenza andrebbe al di là di una valutazione basata su una logica algoritmica che, per ipotesi, si limitasse ad

⁶ Su questo tema vedi M. Ricca, *Culture interdette. Modernità, migrazioni e diritto interculturale*, Torino, 2013, 201 ss.

⁷ Cfr. S. Quintarelli, *op. cit.*, pos. 2023.

⁸ twitter.com/MassimoFaggioli/status/1243616735368024070.

⁹ Si vedano i "Community standards" della piattaforma Facebook, relativamente alla nudità presente nei contenuti degli utenti: www.facebook.com/communitystandards/adult_nudity_sexual_activity.

¹⁰ Cfr. Cambridge Consultants, *Use of AI in online content moderation. 2019 report*, disponibile online: www.ofcom.org.uk/__data/assets/pdf_file/0028/157249/cambridge-consultants-ai-content-moderation.pdf.

una di tipo *true/false*. Appare quindi complesso pensare di risolvere il problema semplicemente istruendo un algoritmo con la regola “Cancella i post riguardanti la *sura* del Corano in cui è presente la parola guerra”, poiché il riferimento alla norma religiosa potrebbe essere in quella sede utilizzata, per esempio, a fini di confutazione di una fake news sull’argomento. Occorre immaginare una sensibilità della macchina particolarmente evoluta per evitare che contenuti legati al radicalismo religioso siano confusi con altri di pura discussione, critica o riguardanti l’esercizio del culto in ambienti digitali.

La pervasività delle norme prodotte dalle religioni fa sì che queste diventino delle vere e proprie abitudini di vita per i fedeli. Tali regole non sempre si presentano in modo univoco, ma subiscono l’influsso del tempo e del cambiamento della tradizione, dell’intreccio con le regole culturali e con le regole giuridiche di una certa comunità¹¹. Del resto, il tasso di pluralismo interno alle religioni, fa sì che a volte la medesima regola sia interpretata in modo diverso¹².

Se persino la Corte di cassazione in un’occasione ha interpretato il tema del porto di simboli religiosi – nello specifico, il *kirpan*, il pugnale sacro dei Sikh – facendone una questione migratoria e di “conflitto di valori”¹³, è lecito aspettarsi che la macchina possa avere difficoltà nell’identificare il *kirpan* come un simbolo religioso piuttosto che come una possibile arma. Per questo un ostacolo propedeutico appare quello di individuare gli strumenti adatti per istruire le macchine sulla comprensione delle implicazioni semantiche della religione, evitando rappresentazioni semplificate, se non semplicistiche. A ben vedere, si tratta di evitare di far ripetere alle macchine lo stesso errore del giudice, che non sempre riesce ad avere un approccio interculturale alla diversità culturale e religiosa, facendo prevalere termini pregiudizievole e stereotipici, che possono condurre a decisioni lesive dei diritti delle persone¹⁴.

3. Prospettive di una “evangelizzazione” dell’IA

La difficoltà per le macchine di interpretare il dato religioso all’interno dei contenuti online può quindi costituire un esempio della necessità di migliorare questo tipo di tecnologie tenendo conto della particolare natura del fenomeno religioso, al fine di predisporre contromisure migliori rispetto agli episodi di *hate speech* presenti nello spazio digitale. In sostanza, per un giudice algoritmo potrebbe essere utile avviarsi ad un’effettiva conoscenza della religione, al fine di poter per lo meno far comprendere una parte della sensibilità umana agli aspetti spirituali. Si potrebbe parlare, mutuando una definizione nata nel mondo delle religioni, di

¹¹ Cfr. M. Ricca, *Pantheon. Agenda della laicità interculturale*, Palermo, 2013, 29 ss.

¹² Vedi S. Ferrari, *Strumenti e percorsi di diritto comparato delle religioni*, Bologna, 2020.

¹³ Il riferimento è a Cass. pen. n. 24084 del 2017. Vedi M. Del Vecchio, *Lo straniero errante. Il kirpan e le ragioni dell’«errore» dei Sikh*, in *Calumet. Intercultural Law and Humanities Review*, n.7, 2018, 77-106.

¹⁴ Cfr. I. Ruggiu, *Il giudice antropologo. Costituzione e tecniche di composizione dei conflitti multiculturali*, Milano, 2012; P. Consorti, *Conflitti, mediazione e diritto interculturale*, Pisa, 2013.

“evangelizzazione” delle macchine, di avvicinamento della mente artificiale alla comprensione di alcuni aspetti della spiritualità umana.

In effetti, di recente l’operato dei meccanismi di intelligenza artificiale ha iniziato a suscitare l’interesse delle stesse comunità religiose, sia per quanto attiene alla tutela della propria libertà religiosa sia per aspetti più generali legati alla giustizia sociale. L’impatto delle tecnologie di intelligenza artificiale nella società, dal punto di vista della Chiesa cattolica, si pone nei termini di un avvicinamento dell’intelligenza artificiale al contenuto valoriale del proprio messaggio religioso. Come è stato osservato riguardo l’impatto sociale delle tecnologie di IA, anche la Chiesa “ha l’opportunità e l’obbligo di impegnare il suo insegnamento, la sua voce e la sua autorevolezza riguardo ad alcune questioni che si profilano fondamentali per il futuro”¹⁵. L’idea è che questa “evangelizzazione” dell’IA debba concretizzarsi nell’istruire le macchine non solamente nei contenuti dogmatici di una particolare tradizione religiosa, bensì nel fornirgli l’accesso a quel bagaglio di esperienza delle religioni in termini di riflessioni sull’etica¹⁶ e sulla giustizia sociale finalizzata alla tutela e al rispetto della persona umana, che una macchina che ragionasse semplicemente in termini razionali non potrebbe sviluppare da sola. In questa prospettiva, le comunità religiose appaiono come interlocutori qualificati nel dibattito sullo sviluppo e sulla gestione di questo tipo di tecnologie¹⁷.

Istruire l’IA sull’esperienza religiosa implica quindi di andare al di là della mera concezione della religione come una questione di appartenenza ad un gruppo¹⁸. Piuttosto, si tratta di inquadrarne i principi religiosi in termini di utilità sociale, mettendo a frutto il contributo della religione al “progresso materiale e spirituale della società”.

Simone Baldetti
Dip.to di Giurisprudenza
Università di Pisa
simone.baldetti@phd.unipi.it

¹⁵ Cfr. A. Spadaro, T. Twomey, *Intelligenza artificiale e giustizia sociale. Una sfida per la Chiesa*, in *La Civiltà Cattolica*, Quaderno n. 4070, Volume I, gennaio 2020, 130.

¹⁶ Sulla necessità di vigilare su di una “etica delle IA” vedi T. Hagendorff, *The Ethics of AI Ethics. An Evaluation of Guidelines*, in *Minds & Machines*, 2020, 1-22.

¹⁷ Cfr. *Response Document To The Public Consultation On The Eu Commission White Paper On Artificial Intelligence*, (2020), elaborato dal Center for Religious Studies della Fondazione Bruno Kessler: “In accordance with its mission and its Strategic Plan 2019-2021, the Center for Religious Studies has contributed to the EU-wide debate on the European Commission’s White Paper “On Artificial Intelligence – A European Approach to Excellence and Trust”. Drawing upon our work on the interaction between religion and innovation, our contribution to the public consultation launched by the Commission in February 2020 contains our Response Paper along with our answers to the consultation questionnaire”. Il documento è consultabile online all’indirizzo web isr.fbk.eu/en/about-us/response-to-the-european-commissions-public-consultation-on-the-white-paper-on-artificial-intelligence/ (ult. visit. 17.07.2020).

¹⁸ Cfr. AA. VV., *The Meaning of ‘Religion’ in multicultural societies law*, in *Stato, Chiesa e pluralismo confessionale*, rivista telematica (www.statoechiese.it), n. 39 del 2017.