

LA COSTOLA MALANDATA.
FILTRI ALGORITMICI E PREGIUDIZI DI GENERE¹

di Paolo Costa

SOMMARIO: 1. Introduzione – 2. I gender script degli artefatti tecnologici – 3. Demistificare l’AI – 4. La discriminazione nel ML – 5. Possibili soluzioni al problema

1. *Introduzione*

La riflessione che qui propongo si articola in quattro parti. La prima parte ha lo scopo di ricordare due condizioni di ordine generale, sovente trascurate, che connotano il nostro rapporto con la tecnologia: da un lato gli artefatti tecnologici non sono neutrali, perché fin dalla loro progettazione contengono modelli di rappresentazione dell’utente, e in particolare quelli che nell’ambito degli studi femministi sono noti come *gender script*; dall’altro lato l’effetto di tali modelli non è meccanico e deterministico, perché gli utenti rinegoziano continuamente i significati degli artefatti tecnologici. Ogni tecnologia risulta dunque come il prodotto di un negoziato culturale e politico fra progettisti e gruppi sociali. Nella seconda parte mi concentro sulle tecnologie dell’intelligenza artificiale (*artificial intelligence*, d’ora in avanti AI), con l’obiettivo di evidenziare come la rappresentazione più diffusa di tali tecnologie non corrisponda alla loro effettiva evoluzione, ma tenda piuttosto a sopravvalutarne le capacità. Con il risultato che si genera una dissonanza fra le prestazioni dell’AI e le aspettative del pubblico. Una simile dissonanza predispone a sua volta a una relazione disfunzionale fra le tecnologie medesime e gli utenti. Nella terza parte mostro come, in particolare, questa relazione disfunzionale si manifesti nell’ambito dei sistemi di apprendimento automatico (*machine learning*, d’ora in avanti ML), che incorporano pregiudizi di genere di vario tipo. Nel tentativo di illustrare il problema con qualche esempio, ho scelto l’ambito dei sistemi per il trattamento del

¹ Intervento al convegno *Parità di genere e nuove tecnologie*, Università degli Studi di Pavia, 26 maggio 2021.

linguaggio naturale, che hanno un forte impatto sulla vita di tutti noi. Infine, nella quarta parte accenno ad alcune possibili soluzioni al problema.

Un breve cenno al titolo di questo contributo. Si tratta di un omaggio a *Eva non è ancora nata*, canzone composta nel 1978 da Giorgio Gaber e Sandro Luporini e inserita nell'album *Polli di allevamento*². Così recita infatti la sesta strofa: « Come son giusti i giovani / Insieme alle loro ragazze / Coscienti che questa polemica è superata / Come son più sensibili / Di fronte a una nuova realtà / Le idee sono buone ma la costola è malandata / Eva non è ancora nata».

2. *I gender script degli artefatti tecnologici*

Riguardo al rapporto fra tecnologia e società esistono due posizioni altrettanto diffuse ai giorni nostri. La prima si fonda sul mito della neutralità tecnologica, ossia sull'idea che nessuna tecnologia abbia, in sé, qualcosa di buono o di cattivo. In base a tale visione a essere più o meno corretto, semmai, sarebbe l'uso che facciamo della tecnologia. La seconda posizione, al contrario, postula fra tecnologia e società un rapporto di tipo deterministico: la tecnologia, appunto, *determina* i suoi effetti secondo leggi di tipo meccanico: ciascuna tecnologia si comporta in un certo modo perché quello è il suo modo di comportarsi, un po' come l'acqua bolle a 100 °C e i salmoni risalgono i fiumi, e noi possiamo farci ben poco.

Vi è una certa ingenuità nell'uno come nell'altro punto di vista. Quando – evocando la prima legge di Kranzberg – affermiamo che la tecnologia non è in sé buona o cattiva, ma non per questo è neutrale, ci riferiamo al fatto che gli sviluppi tecnologici hanno in genere conseguenze sociali, economiche e ambientali che vanno al di là dei loro obiettivi originari. D'altra parte, la medesima tecnologia può avere conseguenze diverse, a seconda del contesto e delle circostanze che ne accompagnano la diffusione [KRANZBERG 1986, pp. 545-546]. I fattori estrinseci, ossia non connessi in modo diretto alle caratteristiche degli artefatti tecnologici, sono dunque decisivi per comprendere la diffusione e l'impatto di tali artefatti.

Il punto sembra risiedere nel fatto che gli artefatti tecnologici si inseriscono in una

² Registrazione dell'omonimo spettacolo, effettuata al Teatro Duse di Bologna il 18 ottobre 1978 per l'etichetta Carosello Records di Milano.

rete di relazioni che coinvolge esseri umani e macchine, competenze, interessi e desideri. Ma in che modo l'artefatto opera all'interno di questa rete? In che modo, segnatamente, l'artefatto condiziona le relazioni degli agenti umani – quelli che oggi siamo soliti chiamare *utenti* – con sé stesso e con altri agenti? Secondo Madeleine Akrich [1992] per rispondere a tale domanda è necessario comprendere in che misura l'agente umano assunto nell'ambito della progettazione dell'artefatto corrisponde all'agente reale. Nel momento in cui formula delle ipotesi sul suo utilizzo, chi disegna una nuova tecnologia iscrive nel proprio progetto la visione del mondo in cui essa si collocherà. Il/la designer produce degli *script*. Tali script sono incorporati nel dispositivo tecnico e definiscono gli attori, lo spazio in cui si muovono e il modo in cui interagiscono. Ma non è solo questo: il/la designer fornisce anche una chiave da utilizzare per interpretare gli eventi futuri. L'artefatto tecnologico diventa uno strumento di conoscenza e di controllo sociale, nel senso che stabilisce norme di comportamento:

Nella misura in cui le scelte progettuali assumono la forma di decisioni su ciò che deve essere delegato e a chi, questo significa che l'oggetto tecnico contiene e produce una specifica geografia delle responsabilità [...] L'innovazione tecnologica non solo può portare a definire rapporti inediti fra le persone e le cose. In aggiunta, essa può anche generare e “naturalizzare” nuovi ordini di casualità, e dunque nuove forme di conoscenza del mondo. [AKRICH 1992, p. 207; traduzione mia]

Alcuni *script* iscritti nell'artefatto tecnologico hanno una connotazione di genere (*gender script*). Questo può dipendere da due circostanze. Da un lato succede che l'agente incorporato nel design di una determinata tecnologia rappresenti il modello di un uomo, in termini di requisiti e preferenze, piuttosto che quello di una donna. Spesso ciò è legato al fatto che tale tecnologia è stata progettata proprio da un uomo e dunque si presenta come una tecnologia maschile [OUDSHOOM, ROMMES, STIENSTRA, 2004]. Dall'altro lato, sempre in ragione del fatto che il design tecnologico è stato affidato a uomini, esso incorpora un agente femminile che non trova riscontro nella realtà, ma è piuttosto il frutto di stereotipi e pregiudizi. Cosicché abbiamo a che fare con tecnologie rivolte alle donne o che rappresentano le donne, per le quali i loro

progettisti maschi potrebbero dire ciò che Gustave Flaubert disse a proposito di Madame Bovary: «c'est moi». In passato è stata evidenziata, per esempio, la presenza costante di bias di genere all'interno della rappresentazione del mondo fornita da Wikipedia, la nota enciclopedia online curata da una comunità di utenti molto vasta ma con un profilo demografico a predominanza maschile [WAGNER, GARCIA, JADIDI, STROHMAIER 2015].

Un caso da manuale di tecnologie per donne progettate da uomini è quello delle macchine per il supporto al lavoro domestico, descritto da Ruth Schwartz Cowan in un memorabile studio del 1983. Schwartz Cowan ricostruisce la storia delle tecnologie domestiche e dell'industrializzazione della casa, mettendo in evidenza il peculiare percorso che ha condotto le donne a essere assunte all'interno di uno schema del lavoro industrializzato sostanzialmente diverso da quello destinato agli uomini. Certo, ci sono alcuni elementi in comune: il ricorso a risorse energetiche non umane, la dipendenza da reti e istituzioni sociali, l'alienazione (nel senso di estraniamento psicologica). Tuttavia, contrariamente al lavoro industrializzato degli uomini, il lavoro industrializzato domestico non è svolto a fronte di un salario, si consuma in un ambiente isolato e non richiede alcun tipo di specializzazione. L'industrializzazione della casa – osserva Schwartz Cowan – è stata in parte determinata da processi sui quali le casalinghe hanno potuto esercitare un controllo molto modesto o addirittura nullo [SCHWARTZ COWAN 1983, p. 14].

3. *Demistificare l'AI*

Quanto fin qui indicato a proposito della complessa interazione fra tecnologia, individui e sistemi sociali ci aiuta a inquadrare, io credo, talune disfunzionalità che caratterizzano il ruolo dell'AI oggi e, in particolare, il problema dei bias che connotano gli algoritmi di ML sempre più utilizzati in ogni ambito della nostra vita. Il problema, nella sua essenza, può essere formulato come segue: gli algoritmi sono disegnati da esseri umani, i quali incorporano i loro pregiudizi inconsci all'interno degli algoritmi stessi o – per meglio dire, come vedremo – nei dati con cui gli algoritmi stessi sono addestrati. Dunque non dobbiamo stupirci quando il ML non lavora in modo

corretto: non vi è nulla di più o di meglio, nel codice software, di quanto vi possiamo avere inserito noi esseri umani.

Invece l'AI viene caricata di aspettative incongrue. Come accade più in generale nel caso del software, le tecnologie dell'AI nutrono il nostro immaginario. Per questo oggi ci attendiamo dall'AI molto più di ciò che essa può darci, nel bene come nel male. Abbiamo in mente qualcosa di molto simile al concetto astratto di *general AI*, ma poi ci confrontiamo ogni giorno con un altro tipo di AI, quella che esiste e che funziona, ossia la cosiddetta *narrow AI*. E fingiamo di non capire che essa è niente di più di un metodo matematico per la predizione, intesa come la risposta numerica più attendibile a qualunque domanda sia appunto gestibile in modo numerico. Il che, se ci pensiamo, è moltissimo. Ma, al tempo stesso, pochissimo.

«Scusa, non ho capito» è il tipo di risposta che otteniamo da un assistente digitale cui attribuiamo erroneamente un eccesso di intelligenza, se gli chiediamo di esercitarsi in ciò che non gli compete: per esempio di individuarsi al di là del suo sé fisico, di valutare l'ipotesi della reincarnazione, di interrogarsi sull'esistenza dell'aldilà o di dichiarare la propria situazione emotiva. È il problema della coscienza dell'AI. Sia che si affronti tale problema nella prospettiva del materialismo biologico, ovvero riconoscendo la capacità di coscienza solo agli organismi biologici, sia che si propenda per una visione tecno-ottimista, ossia si avvalori l'ipotesi che anche la coscienza sia un fenomeno computazionale, come tutti i fenomeni che si manifestano nel nostro cervello, ci si troverà costretti a concludere che oggi ci confrontiamo con una forma sofisticata di AI priva di coscienza.

Susan Schneider ha mostrato come sia difficile appurare in modo incontrovertibile se una determinata tecnologia di AI sia dotata di una propria soggettività e dunque sia in grado di esprimere una coscienza di sé, anche svolgendo prove ben più articolate rispetto al fallibilissimo test di Turing, quali l'Artificial Consciousness Test (ACT), il Chip Test o il test ispirato alla Integrated information Theory (IIT) [SCHNEIDER 2019, pp. 46-71]. Anche in futuro, probabilmente, chi progetta AI dovrà confrontarsi con una sorta di zona grigia e applicare un principio precauzionale: non sviluppare un sistema di AI ogni volta che si sospetta che esso possa essere dotato di coscienza.

D'altronde siamo indotti a domandarci se, nell'eventualità di riuscire un giorno a creare una macchina dotata di coscienza, saremo legittimati a venderla e sfruttarla senza perpetrare una qualche forma di schiavitù [ivi, p. 36].

Demistificare l'AI significa riconoscere due cose. La prima è che la capacità di manipolare simboli, di cui gli algoritmi di ML sono dotati, non equivale a comprendere. Questo era un punto già chiaro a John Searle ed evidenziato nel suo famoso argomento della Stanza Cinese [Searle 1980]. La seconda circostanza è che le macchine funzionano e hanno senso solo quando operano insieme a esseri realmente intelligenti, ossia noi umani [BROUSSARD 2019, pp. 38-39].

Anche se parliamo di *apprendimento* quando ci riferiamo ai sistemi di ML, dovremmo essere consapevoli che stiamo usando tale espressione in senso metaforico, poiché un sistema di ML non apprende nello stesso modo in cui apprende un essere umano. La macchina migliora nell'esecuzione di compiti programmati e automatizzati, ma non acquisisce una conoscenza e una capacità di agire nuove. I dizionari definiscono il ML come «la capacità di un computer di apprendere dall'esperienza» (*Oxford English Dictionary*, 2000) o «il ramo dell'AI relativo alla costruzione di programmi che imparano dall'esperienza» (*A Dictionary of Computer Science*, 2016). Sarebbe più corretto affermare che il ML consiste nell'abilità della macchina di riconoscere alcune proprietà all'interno di un insieme di dati (*data set*) e applicare tali proprietà a un nuovo insieme di dati. L'intelligenza della macchina si limita alla capacità di migliorare le proprie prestazioni con riferimento a uno specifico compito – per esempio distinguere l'immagine di una cellula sana da quella di una cellula tumorale – sulla base di uno specifico criterio di misurazione definito da un essere umano [MITCHELL 2006].

5. *La discriminazione nel ML*

Per comprendere in che modo la discriminazione di genere agisce nell'ambito del ML occorre tornare alla definizione delle tecniche di apprendimento automatico fornita poc'anzi: gli algoritmi analizzano un insieme di dati, riconoscono al suo interno alcune proprietà e quindi applicano tali proprietà a un nuovo insieme di dati. In altri termini, gli algoritmi formulano ipotesi sul mondo basandosi sulla rappresentazione

del mondo stesso che ricevono in input. Dove si insinua dunque la discriminazione? Nel *data set* di partenza, il quale fornisce una rappresentazione del mondo che incorpora pregiudizi di genere. Il *gender gap* dell'AI si configura dunque nella sua essenza come un *gender data gap*. Va detto che la questione non è del tutto nuova. Sono numerose le scelte tecnologiche compiute a partire da una rappresentazione parziale del mondo, nella quale il mondo appare abitato solo da uomini, preferibilmente bianchi ed eterosessuali. Il problema, ad esempio, si è posto fino a poco tempo fa – e forse si pone ancora oggi – nel campo della ricerca farmacologica e della terapia medica: nella misura in cui i grandi studi clinici coinvolgono un numero insufficiente di pazienti di sesso femminile, ossia lavorano su un *data set* sbilanciato dal punto di vista del genere, non deve sorprendere come gli indirizzi terapeutici che da tali studi dipendono si rivelino meno efficaci per le donne, con costi sociali enormi [GRIGLIÉ, ROMEO 2021, pp. 11-32].

In conseguenza dell'uso sempre più estensivo del ML il problema sta assumendo proporzioni nuove, perché impatta su servizi universali e di uso quotidiano. Ancora nel 2017, per esempio, l'efficacia della tecnologia di Google per il riconoscimento della voce risultava inferiore del 13% per le utenti donne rispetto agli uomini [TATMAN 2017]. In altri termini, una donna aveva il 13% di probabilità in più di non essere compresa dall'assistente vocale di Google, un servizio popolare per chiunque possieda un dispositivo con sistema operativo Android sul proprio telefono, nel navigatore dell'auto o nell'altoparlante intelligente di casa. Analoghe disparità di genere, connesse a una inferiore accuratezza degli algoritmi, si riscontrano nell'ambito dei sistemi per il riconoscimento facciale, che sono largamente usati – anche se sempre più discussi – sia nell'ambito della videosorveglianza sia in quello del marketing [BUOLAMWINI, GEBRU 2018].

Ovviamente l'argomento si presta a essere ribaltato. Da un lato le donne sono discriminate nell'uso di determinate tecnologie, dal momento che tali tecnologie sono progettate per funzionare meglio con gli uomini. Dall'altro lato esistono tecnologie progettate per funzionare solo con le donne, delle quali però sospetto che le donne stesse farebbero volentieri a meno. Come valutare, per esempio, un algoritmo

progettato per calcolare – con l’analisi di specifiche tracce, lasciate inconsapevolmente online attraverso applicazioni personali per la gestione del calendario mestruale, come Maya o MIA Fem – le probabilità che una donna sia incinta o addirittura per stabilire in quale fase del ciclo mestruale ella si trovi, per poi sottoporle proposte commerciali personalizzate? Il passaggio di dati da queste applicazioni a Facebook è stato documentato da Privacy International in uno studio pubblicato nel 2018, il quale ha avuto se non altro il merito di costringere Maya, MIA Fem e la stessa Facebook a interrompere questa forma di sorveglianza intima (rimando al documento *No Body’s Business But Mine: How Menstruation Apps Are Sharing Your Data*, disponibile sul sito web di Privacy International: <https://www.privacyinternational.org/long-read/3196/no-bodys-business-mine-how-menstruations-apps-are-sharing-your-data>).

Il caso d’uso più interessante dal mio punto di vista riguarda però le tecniche di AI che si sforzano di inseguire le competenze semantiche dell’essere umano, ossia di comprendere il significato delle espressioni formulate nei linguaggi naturali. Robert (Munro) Monarch e Alex (Carmen) Morrison hanno dimostrato che ancora alla fine del 2019 tecnologie di NLP (Natural Language Processing) molto popolari, come Amazon Comprehend, Google Natural Language API e Stanford Parser, non riescono a riconoscere l’espressione inglese *hers* come un pronome femminile [MONARCH, MORRISON 2020]. In una frase come «the car is hers», ossia «l’auto è sua (di lei)», Amazon e Google classificano *hers* come nome, Stanford Parser come aggettivo. Anche BERT (Bidirectional Encoder Representations), l’algoritmo introdotto da Google proprio nel 2019, che potenzia molto le capacità semantiche del motore di ricerca, sembra affetto dallo stesso problema. Il quale tuttavia non si manifesta con il corrispondente pronome maschile *his*, a conferma del fatto che esiste un pregiudizio di genere iscritto in tutte queste tecnologie. I due autori formulano diverse ipotesi in merito alle possibili cause di questa evidente disfunzionalità. Il problema potrebbe essere connesso all’addestramento svolto su *data set* squilibrati dal punto di vista della diversità di genere o poveri di informazioni utili (per esempio gli articoli di giornale, in cui prevalgono le forme non connotate dal punto di vista del genere), oppure al mancato coinvolgimento di esperti di dominio (ovvero linguisti).

Vengono invece escluse altre ipotesi: la presenza di bias intrinseci al linguaggio (in questo caso la grammatica inglese) o agli algoritmi, oppure l'esistenza di pregiudizi inconsci nella cultura di chi ha programmato tali algoritmi. In sostanza, il difetto non starebbe negli algoritmi né nei loro autori, ma nel modo in cui gli algoritmi stessi sono stati addestrati. A risultati non dissimili sono pervenuti anche altri studi, più o meno recenti [BOLUKBASI ET AL. 2016].

5. Possibili soluzioni al problema

Proprio il fatto che il problema sembra risiedere principalmente nella metodologia di addestramento e quindi nei *data set* impiegati a tale scopo suggerisce quale strada intraprendere per cercare una soluzione. Si tratta, in sostanza, di identificare i bias presenti nei dati di training, anche nel caso in cui essi sono più difficili da riconoscere. Impresa particolarmente complessa, ovviamente, quando si tratta di sistemi di NLP. D'altra parte il linguaggio ha un ruolo fondamentale, come sappiamo, nella costruzione e nel rafforzamento di una prospettiva di genere. Pensiamo a espressioni come *donna in carriera* o *ragazza madre*, che non trovano un corrispettivo maschile nella lingua italiana e sono entrambe connotate in senso tendenzialmente negativo.

Un approccio interessante – ancorché suscettibile, con ogni probabilità, di provocare profonde discussioni – è stato formulato recentemente da Susan Leavy et al. Esso si fonda sull'applicazione della linguistica femminista e degli studi di genere per snidare i bias all'interno di un *data set*. Il metodo introduce una serie di metriche relative alla presenza proporzionale, nel corpus di training, di nomi/pronomi maschili e femminili, così come di termini, professioni, caratteristiche (per esempio *modestia*, *potere*, *fragilità*) e riferimenti anatomici connotati [LEAVY, MEANEY, WADE, GREENE 2020].

Non sappiamo se questa o altre proposte, animate dallo stesso obiettivo, possano contribuire in concreto a risolvere un problema che ha anche risvolti industriali, dato che il dominio sulle tecnologie di ML è esercitato da uno sparuto gruppo di *provider* globali, i quali difficilmente adotteranno nei propri processi metodologie non scalabili. Quello che è certo è che la strada da percorrere è ancora lunga. Come cantava Gaber,

le idee sono buone ma la costola è malandata.

Abstract

Our relationship with technology reflects two general conditions: on one side, technological artifacts are not neutral. Their design contains user representations, and in particular those models that in feminist studies are known as gender scripts; on the other hand, the effect of these models is not mechanical and deterministic, because users continuously negotiate the meanings of technological artifacts. Each technology thus turns out to be the product of a cultural and political negotiation between designers and social groups. Focusing on artificial intelligence, we can see how the most widespread representation of these technologies does not correspond to their actual evolution, but rather tends to overestimate their capabilities. The result is a dissonance between AI performance and public expectations. This dissonance in turn predisposes to a dysfunctional relationship between the technologies themselves and the users. This dysfunctional relationship is evident in the context of machine learning systems, which incorporate gender biases of various kinds. To illustrate the problem with a few examples, I have chosen the domain of natural language processing systems, which have a strong impact on the lives of all of us. In the last part of this paper I mention some possible solutions to the problem.

Bibliografia

Madeline Akrich, *The De-Description of Technical Objects*, in Wiebe E. Bijker, John Law (a cura di), *Shaping Technology / Building Society: Studies in Sociotechnical Change*, Cambridge MA, The MIT Press, 1992, pp. 205-224.

Tolga Bolukbasi, Kai-Wei Chang, James Zou, Venkatesh Saligrama, Adam Kalai, *Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings*, in Proceedings of the 30th International Conference on Neural Information Processing Systems (dicembre 2016), pp. 4356–4364.

Joy Buolamwini, Timmit Gebru, *Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*, in *Proceedings of Machine Learning Research. Conference on Fairness, Accountability, and Transparency*, 81 (2018), pp. 1-15.

Meredith Broussard, *Artificial Unintelligence. How Computers Misunderstand the World*, Cambridge MA - London, The MIT Press, 2019.

Emanuela Griglié, Guido Romeo, *Per soli uomini. Il maschilismo dei dati, dalla ricerca scientifica al design*, Torino, Codice, 2021.

Melvin Kranzberg, *Technology and History: "Kranzberg's Laws"*, in «Technology and Culture», 27, 3 (luglio 1986), pp. 544-560, DOI: <https://doi.org/10.2307/3105385>.

Susan Leavy, Gerardine Meaney, Karen Wade, Derek Greene, *Mitigating Gender Bias in Machine*

Learning Data Sets, in Ludovico Boratto, Stefano Faralli, Mirko Marras, Giovanni Stilo (a cura di), *Bias and Social Aspects in Search and Recommendation. BLAS 2020. Communications in Computer and Information Science*, Cham, Springer, 2020, pp. 12-26, DOI: https://doi.org/10.1007/978-3-030-52485-2_2.

Tom M. Mitchell, *The Discipline of Machine Learning*, Pittsburg PA, School of Computer Science, Carnegie Mellon University. 2006: <http://www.cs.cmu.edu/~tom/pubs/MachineLearning.pdf>.

Robert (Munro) Monarch, Alex (Carmen) Morrison, *Detecting Independent Pronoun Bias with Partially-Synthetic Data Generation*, in *The 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*: http://robertmunro.com/research/MonarchMorrison2020IDP_Bias.pdf.

Nelly Oudshoorn, Els Rommes, Marcelle Stienstra, *Configuring the User as Everybody: Gender and Design Cultures in Information and Communication Technologies*, in «Science, Technology, & Human Values», 29, 1 (gennaio 2004), pp. 30-63, DOI: <https://doi.org/10.1177%2F0162243903259190>.

Susan Schneider, *Artificial You. AI and the Future of Your Mind*, Princeton NJ - Oxford, Princeton University Press, 2019.

Ruth Schwartz Cowan, *More Work for Mother: The ironies of household technology from the open hearth to the microwave*, New York NY, Basic Books, 1983.

John Searle, *Minds, Brains and Programs*, in «Behavioral and Brain Sciences», 31 (1980), pp. 417-457.

Rachel Tatman, *Gender and Dialect Bias in YouTube's Automatic Captions*, in *Proceedings of the First Workshop on Ethics in Natural Language Processing*, Valencia (4 aprile 2017), pp. 53-59: <http://www.ethicsinnlp.org/workshop/pdf/EthNLP06.pdf>.

Claudia Wagner, David Garcia, Mohsen Jadidi, Markus Strohmaier, *It's a Man's Wikipedia? Assessing Gender Inequality in an Online Encyclopedia*, in *The International AAAI Conference on Web and Social Media (ICWSM2015)*, Oxford, 2015.