



Nexa Center for Internet & Society

Politecnico di Torino

Per il ciclo di incontri "i Nexa Lunch Seminar" (ogni 4° mercoledì del mese)

81° Nexa Lunch Seminar

From equality to equity: an interdisciplinary perspective on mitigating algorithmic unfairness in ranking systems

Elena Beretta (Centro Nexa)

Mercoledì 23 settembre 2020 ore 13 in punto

(termine: ore 14)

**Centro Nexa su Internet e Società
Politecnico di Torino, Via Boggio 65/a, Torino (1° piano)**

**L'incontro si svolgerà in modalità telematica utilizzando il software di conference call
JITSI MEET**

Per partecipare all'incontro: <https://meet.nexacenter.org/LunchSeminar>

Live streaming: <https://nexa.polito.it/upcoming-events>

URL dell'incontro: <https://nexa.polito.it/lunch-81>

Nell'ultimo decennio abbiamo assistito a una cospicua diffusione di modelli predittivi per il processo decisionale progressivamente più sofisticati, che sfruttano una quantità sempre maggiore di **dati personali** suggerendo come i dati stessi debbano essere interpretati e quali azioni debbano essere perseguite. Le **decisioni algoritmiche** sono il risultato dei massicci e sempre più potenti **meccanismi di profilazione e classificazione** utilizzati



nell'**apprendimento automatico**, che si sta gradatamente e pervasivamente diffondendo nelle nostre vite. In molti campi in cui questi sistemi sono stati applicati i progressi sono stati significativi. Tuttavia, una quantità rilevante di studi dimostra che, come gli esseri umani, gli **algoritmi** sono **suscettibili di pregiudizi**. A questo proposito, negli ultimi anni, ricercatori e professionisti hanno portato alla luce una

<http://nexa.polito.it/contact>

serie di problematiche legate ai **risultati discriminatori di alcuni processi decisionali automatizzati**, in particolare nei confronti dei gruppi svantaggiati. Conseguentemente, le comunità di ricerca stanno dedicando considerevoli sforzi per studiare come includere il concetto di equità negli algoritmi di apprendimento automatico al fine di produrre risultati più equi ed evitare risultati discriminatori. Le attuali ricerche formalizzano il concetto di equità tramite diversi approcci, mostrando così una vasta e diversificata gamma di soluzioni sviluppate per essere applicate in contesti specifici. Nonostante tali soluzioni siano nate per regolare lo sviluppo, la diffusione e la governance responsabile dei sistemi di apprendimento automatico, la **ricerca sull'equità algoritmica** è in realtà parte di un più ampio dibattito su come queste tecnologie incorporino implicitamente alcune visioni politiche, economiche o culturali. Nonostante diversi sforzi siano stati compiuti in questa direzione, le questioni aperte sono ancora molte, in particolar modo per il fatto che la ricerca in questo campo si è spesso concentrata sul fornire una definizione di equità, piuttosto che cercare delle soluzioni alla disuguaglianza di fondo. Durante il corso del seminario verrà illustrato come strumenti e metodi dell'**informatica** e delle **scienze sociali** possano essere combinati **per definire, misurare e mitigare tale iniquità algoritmica**. In particolare, verrà illustrato **AFTeRS**, un **sistema di ranking automatizzato equo-distributivo** per la giustizia sociale nel campo dei sistemi automatici, che affonda le sue radici nelle teorie filosofico distributive e in quelle economiche di uguaglianza delle opportunità.

Biografia:

Elena BERETTA received the M.Sc. in Economics and Statistics from University of Turin in September 2016 by working on an experimental thesis investigating the diffusion of innovation by agent-based models. She earned a second level Master degree in Data Science for Complex Economic Systems at the Collegio Carlo Alberto in Moncalieri (TO), in June 2017. From April 2017 to September 2017 she got involved in an internship at DESPINA - Laboratory on Big Data Analytics at the Department of Economics and Statistics of the University Study of Turin – working on the NoVELOG project ("New Cooperative Business Models and Guidance for Sustainable City Logistics"). In November 2017 she's starting to collaborate as PhD student, and effective member, with the Nexa Center for Internet & Society at Politecnico of Turin and with the Fondazione Bruno Kessler (Trento), by working on a project on Data and Algorithms Ethics. Her current research focuses on improving the impact of automatic decision-making systems on society through the implementation of models involving

data on human behavior. Specific fields of interest include data science, machine learning, recommendation systems and computational social sciences.

Lecture consigliate e link utili:

- Barocas, S., Hardt, M., Narayanan, A. (2018). Fairness and Machine Learning. <http://www.fairmlbook.org>
- Benjamin, R. (2019). Assessing risk, automating racism. *Science*, 366(6464), 421–422.
- Gummadi, K. P., Heidari, H. (2019). Economic Theories of Distributive Justice for Fair Machine Learning. In *Companion Proceedings of the 2019 World Wide Web Conference, WWW '19*, pp. 1301–1302. ACM.
- Roemer, J. E., Trannoy, A. (2015). Equality of opportunity. *Handbook of Income Distribution*, 2(2), 217–300.
- Singh, A., Joachims, T. (2018). Fairness of exposure in rankings. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD'18*, p. 2219–2228, New York, NY, USA. Association for Computing Machinery.

Che cosa sono il Centro Nexa e i cicli di incontri "Mercoledì di Nexa" e "Nexa Lunch Seminar"

Il Centro Nexa su Internet & Società del Politecnico di Torino (Dipartimento di Automatica e Informatica) dal 2006 è un centro di ricerca indipendente e interdisciplinare che, in collaborazione con l'Università di Torino, studia Internet (e più in generale le Tecnologie digitali) i suoi rapporti con la società. Maggiori informazioni all'indirizzo: <http://nexa.polito.it/about>.

Durante i "Mercoledì di Nexa", che si tengono **ogni 2° mercoledì** del mese alle **ore 18 in punto**, il Centro Nexa su Internet e Società apre le sue porte non solo agli esperti e a tutti coloro i quali lavorano con Internet, ma anche a semplici appassionati e cittadini. Il ciclo di incontri intende approfondire, con un linguaggio preciso ma accessibile, i temi legati alla Rete: motori di ricerca, Creative Commons, social networks, open source/software libero, neutralità della rete, libertà di espressione, privacy, file sharing, big e open data, smart cities e molto altro.

Al centro di quasi tutti gli incontri un ospite pronto a dialogare con i direttori del Centro Nexa, il Prof. Juan Carlos De Martin del Politecnico di Torino e il Prof. Marco Ricolfi dell'Università di Torino, nonché lo staff e i Fellows del Centro Nexa.

Maggiori informazioni sui Mercoledì di Nexa, incluso un elenco di tutti i "Mercoledì" passati, sono disponibili all'indirizzo: <http://nexa.polito.it/mercoledi>.

Si segnala inoltre che dal maggio 2012 **ogni 4° mercoledì** del mese **dalle ore 13 alle ore 14** il Centro Nexa organizza anche i "**Nexa Lunch Seminar**". Una lista di tutti i "Lunch Seminar" passati è disponibile all'indirizzo: <http://nexa.polito.it/lunch-seminars>.

Per restare aggiornati sulle attività del Centro Nexa su Internet & Società, seguitemi su:

- Twitter: [@nexacenter](https://twitter.com/nexacenter)
- Facebook: [http://facebook.com/nexa.center](https://facebook.com/nexa.center)
- Mailing list degli annunci Nexa: <https://server-nexa.polito.it/cgi-bin/mailman/listinfo/nexa-announc>

*Responsabile Comunicazione Centro Nexa su Internet & Società: **Anita Botta**, tel: +39 011 090 7219, Mob: +39 347 131 3903, anita.botta@polito.it*

<http://nexa.polito.it/contact>